2004-06

Identification and resolution of problems with methodology used in selection of technological concepts for R & D support

Hardman, William L.

Monterey California. Naval Postgraduate School

Identification and resolution of problems with methodology used in selection of technological concepts for R&D support.

by

William L. Hardman

B.S. Geophysics, University of Utah, 1991
B.S. Physics, University of Utah, 1992
MBA, University of Utah, 2000

Submitted to the Departments of Ocean Engineering and Nuclear Engineering in Partial Fulfillment of the Requirements for the Degrees of

Naval Engineer
and
Master of Science in Nuclear Engineering

at the
Massachusetts Institute of Technology

June 2004

Author ......William L Hardman......
Department of Ocean Engineering
July 16, 2003

Certified by ......Michael W. Golay......
Michael W. Golay
Professor of Nuclear Engineering
Thesis Supervisor

Certified by ......Henry S. Marcus......
Henry S. Marcus
Professor of Marine Systems
Thesis Reader

Accepted by ......Jeffrey A. Coderre......
Jeffrey Coderre
Associate Professor of Nuclear Engineering
Chairman, Departmental Committee on Graduate Studies

Accepted by ....
Michael Triantafyllou
Professor of Ocean Engineering
Chairman, Departmental Committee on Graduate Studies

Identification and resolution of problems with methodology used in selection of technological concepts for R&D support.

by

William L. Hardman

Submitted to the Departments of Ocean Engineering and

Nuclear Engineering

on July16,2003 in partial fulfillment of the

requirements for the degrees of

Naval Engineer

and

Master of Science in Nuclear Engineering

Abstract

How to better evaluate the merits of competing technological concepts.
This project is concerned with evaluation of the merits of competing technological concepts and managing the R&D investments needed to bring them to fruition. A weakness of the current R&D process is that arguments regarding who should receive funding come from a concept's proponents, who are usually interested and biased. More objective evaluation methods are needed. As part of this work it is important to understand better how qualified experts evaluate technological concepts. Can a probabilistically formulated method of integrating knowledge of various performance attributes provide better understanding of the likely performance of a technological concept? This is the question of interest.

A nuclear power plant example (impetus for the actual study below).
The impetus for the study began with The U.S. Department of Energy's Generation IV advanced reactor technology program, the program that will select the next generation of nuclear reactors. Generation IV chose twenty-seven criteria for use in determining which nuclear power plant concepts would be best for a given mission. These criteria came in the form of twenty-seven questions asked of prospective concept designers. The concept designers ranked their own design over a range of seven bins and specified a peak in the most likely bin. The 27 criteria were assumed to be independent and were used in creating three major goals (sustainability, safety & reliability, and economics). That is, the score assigned in each of the 27 areas was rolled into 3 major scores called goals in this study. Weights, unknown to the concept designers, were assigned to individual

questions and the three major goals, and then probability mass functions were created predicting the success of a given design.

A robot design course.
At the Massachusetts Institute of Technology (MIT) between Fall semester and Spring semester (i.e., during January) an independent activities period (IAP) offers the opportunity for students to design robots in MASlab (Mobile Autonomous Systems Laboratory) Robotics Competition, also known as course 6.186. Therefore, course 6.186 provides an opportunity for evaluating technological concepts (i.e., in the form of a robot design as well as operational contests of those designs). Course 6.186 provides an opportunity for students to act as consultants in offering their expertise in the evaluation of robots designed by themselves and their competitors. The evaluations are composed of questionnaires similar to those described in the nuclear power example. The consultants' responsibilities are to evaluate the strengths and weaknesses of other robots. The quality of the evaluations is indicated by the results of the robot competition.

From this experience we learn more about how objective evaluations of the performance of competing concepts can be made. As coordinators of this effort, we identify the methodology of those consultants who were most successful in identifying, before testing, the best robot designs. The methodologies thus identified can be extended to large-scale projects in general such as identification of the best, among competing, technological concepts.

Thesis Supervisor: Michael W. Golay
Title: Professor of Nuclear Engineering

Thesis Reader: Henry S. Marcus
Title: Professor of Marine Systems

Acknowledgements

I would like to thank Prof. Michael Golay for not only being a very supportive advisor and providing the impetus for the thesis, but also for giving much needed and timely guidance throughout the analysis and writing processes.

I would like to thank Mr. Edwin Olson and the other staff members of the Mobile Autonomous Systems Laboratory (Maslab), as well as the students participating in Maslab for allowing me to collect the necessary data even though their own time was highly constrained in running and preparing for the Maslab contest, staff and students respectively.

This thesis is dedicated to:

My wife: Michele J. Hardman
My parents: William W. Hardman and Helma S. Hardman

Contents

List of Figures

List of Tables

List of Equations

Chapter 1

Introduction

1.1 Literature search

Two examples were covered in the literature search, which follows below. The two examples are the Navy DD (X) acquisition process and how technology investments are targeted in the drug delivery process. Both examples provide insight into the methodology used in selection of technological concepts for R&D support, and, therefore, were deemed useful to our project.

1.1.1 Navy DD (X) acquisition process example [1]

Because of the large expense in designing and producing a naval vessel (i.e., on the order of billions of dollars), the R & D process leading to the detailed design and construction is unique to this industry. First, the two major corporations owning shipyards in the United States, Northrop Grumman and General Dynamics, created an alliance to ensure that both shipyards would participate in the construction of the DD (X), thus ensuring the competitive capability of both corporations' shipyards. A competition followed between the two corporations, called teams, to obtain what the Navy calls the DD (X) design agent contract which includes the design of the ship and building and testing of engineering development models (i.e., prototypes of new systems).

The above process is part of a five-phase process of which the first two phases are described in this paragraph. Phase I involved both teams in ship design and support concepts, performance specification, and system and support requirement determination.

Phase II involved both teams in preliminary ship and support design and initial system specification.

Phase III is the point at which down-selection to one design occurs, which occurred on 29 April 2002 when Northrop Grumman became the design agent. Phase III is when the engineering development models will be built and tested. Phase III also includes production readiness, complete system and support design, and critical design. It is worth noting again, that although one team's design was selected, this does not exclude the other team. "Both design teams come together as a single team and participate in the design of the ship, so that both yards are prepared to build the ship when it comes time to have the competition in 2005. "

The competition spoken of above is that for the lead ship contract and leads to Phase IV. "We expect to have a second competition or the next-step competition to award detailed design and final construction of the lead ship." Again, both teams will participate once the competition has been decided except that the "lead designer" will, by definition, have the lead seat at the table in making decisions about the design. Finally, Phase V occurs which is simply the execution of engineering and logistics life cycle support of the DD (X).

1.1.2 Targeting technology investments in the drug delivery process [2]

The acquisition process for a naval vessel is described above, offering insight into the steps the Navy must follow to target its technology investments (i.e., the selection of new vessel designs and the subsequent construction of these new vessels).

Pharmaceutical companies have similar hurdles in targeting their own technolgy investments in the drug delivery process.

In 1990, the field of genomics (the study of DNA) was created. "Comparing DNA sequences from a number of species enables scientists to study the function of a gene as it is expressed in different organisms. Similarities in gene function between these organisms and humans enable pharmaceutical researchers to use these organisms to analyze the effect of new therapeutic agents on the biological function of specific proteins, i.e., target molecules." The huge volume of genomic data made the creation of a new field known as bioinformatics essential. Bioinformatics is used to extract genomic data from public and private databases via programs written to better understand biological processes. "Bioinformatics, linked to other new technologies, combinatorial chemistry and high throughput screening, has created a new paradigm for drug discovery. This technological change will move the industry from serendipitous discovery of new drugs to strategic management of markets and technology to improve healthcare for targeted diseases."

"Technology management is critical to the pharmaceutical industry for a number of reasons. First, increased investment in pharmaceutical R&D has not resulted in a significant change in the number of new pharmaceutical agents introduced into the market. The cycle time for development of new drugs and R&D dollars spent per product has increased. This lack of R&D productivity is caused by many problems but among them are understanding the limitations of new drug discovery technologies (e.g., combinatorial chemistry, assays to validate targets), the unpredictable complexity of

biological systems, and difficulty in making appropriate investments in technological areas that increase research productivity."

"Second, many of the biotech companies that provide services to the pharmaceutical industry have moved to internal development of their own drugs either alone or in partnership with established companies. In a recent survey, it was estimated that approximately 50 percent of the drugs sold by the pharmaceutical industry were licensed in from other firms."

To increase research productivity at the corporate level, a process known as roadmapping can be applied. Roadmapping is a technology management tool that can be used for "assessing the potential value of new technologies in meeting the challenges of the drug discovery process." Roadmapping ensures that the right technological capabilities are in place to obtain the desired result. This presupposes that the technology is not only available but also aligned to meet the desired need. Of note is the fact that a roadmap, on the corporate level, focuses on "improving internal processes which may need improvement to increase R&D productivity or to upgrade a step in the drug discovery process that has fallen behind 'industry standards.' The roadmapping process can be divided into five steps: Team Formation, Focus, Technology/Workflow Analysis, Implementation, and Review."

Team formation ensures the inclusion of the appropriate people on the team: people from R&D and technology management, from business development, from finance, from medicinal chemistry, high throughput screening, regulatory, and safety studies. Aside from inviting the appropriate people, "the first priority is to establish a common understanding of the process and the terminology employed in the analysis.

Taking this step will minimize the confusion that can potentially arise during the roadmapping process."

The focus step is where the team begins the development of a "detailed analysis of the drug discovery process." In this analysis technology is an important part. The introduction of new technology is valued provided the process is, thus, made more efficient and effective. "One approach to develop a systematic analysis is to apply Goldratt's Theory of Constraints" in identifying constraints and weaknesses.

"In the drug discovery process, application of this theory must be modified based on limitation of technology and knowledge of biological systems. In the process of performing this analysis, the team must decide the metrics/ factors required to evaluate each step in the process. Issues to consider are costs, predictability of outcome, internal competencies in the organization, and opportunities for technology improvement. Upon completion of the analysis, a model can be created to identify steps in the process that have the greatest potential impact on increasing research productivity."

The input to the technology/workflow analysis is the deliverable from the focus step. Where the deliverable from the focus step is the "identification of a specific step in the process for improvement based on technology availability and the probability of obtaining a successful outcome. The value of technologies in the specific process step is dependent on the degree of their alignment with the needs of the process. Rigorous evaluation of this alignment must be performed to understand the limitations and benefits of the technology."

A useful tool in employing the appropriate technology is to use a matrix with the process needs on the left side of the matrix and potential technologies listed at the top.

The technologies will be ranked either high, medium or low depending on how well aligned the technologies are with respect to the process needs.

Following selection of the appropriate technologies, implementation requirements must be identified and forwarded to management. Management ensures that the appropriate resources, both budgetary and personnel, are made available. "One of the benefits of the roadmapping process is the higher probability that implementation goes according to plan, since multiple functional areas were involved in the formulation of the roadmap and provided expert input to ensure its success."

The review step then follows where "any systematic improvement to a process requires that the team members learn whether the process modification led to the desired outcome and, if appropriate, take further corrective actions. The Deming Cycle (plan-do-check-act) is a useful tool for reviewing the results from any process improvement. The check step permits the review of the agreed-upon metrics to determine whether they have met the goals of the roadmapping effort. If modifications are required, the appropriate members of the team should meet and define a revised action plan. In some cases, this may require modifying an existing technology/process or finding a new technology."

1.2 Background

1.2.1 Nuclear power plant example

The U.S. Department of Energy Generation IV Reactor Program is responsible for selecting and funding initial development of the next generation of nuclear power plants. This program intended to rank its choices for selection based upon scores assigned

quantitatively to each of the power plant concepts. The process chosen used a scoring model.

The scoring process begins with questionnaires filled out by the concept designers for each of the nuclear power plants. Each concept designer answers questions regarding the projected success of the concept with respect to a set of performance goals. The questions asked allowed the designers to select the range of success in the area in question and the potential peak or most likely degree of success. In this study, seven bins (or discrete scores) comprised the entire range from which a score could be selected, with bin one being the lowest score and bin seven the highest.

The Generation IV program assigned probability mass functions based upon the range and peak selected by the designers for each question. For example, if a concept designer selected bins 3,4, and 5 with a peak selected at 4 for a specific question, then the Generation IV might assign a probability mass function of 0.2 for bin 3, 0.6 for bin 4 and 0.2 for bin 5. This same process was followed for each of the questions. Each of the questions fit into one of three major sub-goals. Weightings (with values unknown by the concept designers) were assigned to each of the specific questions allowing probability mass functions to be calculated for each of the three major goals. Weighting were also assigned to each of the three major sub-goals allowing an overall probability mass function to be calculated for each of the nuclear power plant concepts.

1.2.2 Reasons for utilization of data from the robot design contest

Potential problems with the above process exist. First, the work conducted by the Generation IV program was purely hypothetical (i.e., none of the nuclear power plant concepts has actually been built). Secondly, because none of the power plants was built,

this means that none could be compared operationally against their concept designer's projections of success (as shown in the probability mass functions created via the questionnaire). Third, because the concept designers themselves critiqued their own design, inherent bias exists. This is obvious due to the fact that research and development money, followed by the likely prospect of construction of the power plant would be the prize for the victorious concept. Fourth, multiple operational comparisons would give a better idea of the overall operational capability of any one power plant. This process outlined in the steps above is not feasible in the context of prospective nuclear power plants because of the prohibitive cost involved. Other potential problems than those listed above might also exist.

The prospect of using the Mobile Autonomous Systems Laboratory for this research is exciting for all of the reasons that the nuclear power plant data are not. First, multiple robot designs with the same end requirements were actually built. Second, because the robots were built, they could be tested and compared against the initial projections of success by each robot design team. Third, projections of success were not only made by the individual concept designers, but also by their peers (i.e., competitors), allowing for a much less biased evaluation. In addition, none of the competitors had anything to gain monetarily from any particular evaluation ensuring a less biased evaluation by all parties. Fourth, three competitive events occurred in the course of the robot matches (i.e., two exhibition matches followed by the competition) allowing for a more exact assessment of the overall capability of any one robot design.

Chapter 2

Rules of the robot contest

2.1 The playing field

The playing field used in exhibition 1 is shown in Figure 1. The playing fields used in exhibition 2 and in the competition were similar to that shown in Figure 1. The robot teams did not know the shape or size of the playing field prior to any of the three events. The characteristics of the playing field that were known were the following: either six or 12 inch tall white walls with a blue stripe at the top, green floors, red targets, and yellow scoring areas. The fact that the size and shape of the playing field were unknown prior to an event required that the robots be able to respond dynamically to their sensor data [3].

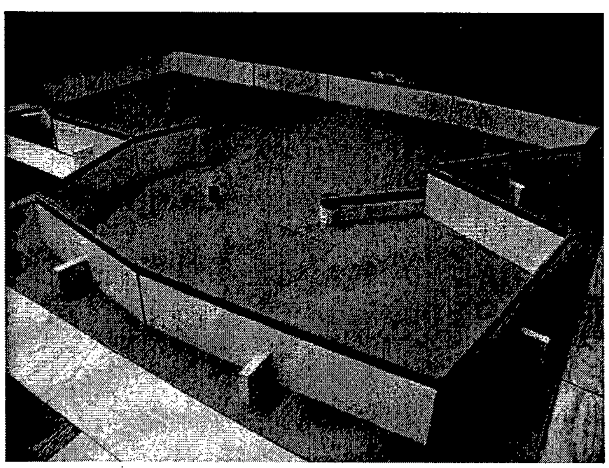

Figure 1: Playing field for exhibition 1 [3]

2.2 How points were awarded

In the contests, four minutes were allowed for each robot in the playing field. Numerous scoring factors were used. See Table 1 below for a complete description of the various scoring factors [3].

Table 1: Methods by which robots score points

Action (of robot)	Point(s)
Approaches target and signals	1
Transfers target to yellow scoring area	3
Transfers target to home	5
Returns to home prior to end of round	1
Moves after four minutes	-1

The first points that were awarded in all cases were for approaching a target, a small red can, and signaling. Signaling meant that the robot recognized the target and made some visible display to that effect. Signaling would have been any recognizable action such as spinning, playing music, making a noise such as a whistle, raising a flag, etc. One point was awarded for each target that was approached and for which a signal was accomplished [3].

After signaling, the robot grasped the target, whether that meant physically attaching the target to the robot by means of a magnetic arm, grasping with pincers, dragging as the robot moved over a target, or simply carrying the target inside or on top of the robot. Taking the target to a yellow scoring area or to home was essential to scoring the next set of points. Transporting the target to within eight inches of a yellow scoring area (yellow rectangular patch on the wall of the playing field) granted the robot

three points for each target left at the yellow scoring area. Five points, however, were awarded if the robot instead transferred a target to home. Home was the location used as the starting point for each of the competing robots at the start of the four-minute playing time [3].

Two final means existed by which points were awarded/deducted. These were either returning home prior to the end of the allotted four minutes (for which one point was awarded) or continuing to move after the four minutes have passed (for which one point was deducted as a penalty) [3].

Chapter 3

Description of data obtained via questionnaires

3.1 Projections of future performance made by participating teams

The questions were divided into three groups called goals. The three goals were labeled as control, control/score, and score. The names of the three goals are illustrative of the content of the questions within the goal grouping. This means that the control group contains questions regarding robot control; the control/score group contains questions, which demonstrate a combination of robot control as well as scoring ability; and finally the score group includes questions, which specifically look at the scoring capability of the robots. The control/score group, also, was broken into two subgroups called category 1 and category 2 which illustrate capabilities relating to targets and scoring respectively.

In the first and second exhibition matches, 17 questions were used to project the future success of each of the teams in the exhibitions (Appendix A). The evaluations provided by the participating teams are found in Appendix B and C for Exhibitions 1 and 2, respectively. The questionnaire for the competition match (i.e., the match following the second exhibition match) included two additional questions (number 4: robot's arrival at targets and numbers 18-23 which ask for the projected success ranking of each of the participating teams). The questionnaire for the competition match also deleted question number 3 from the 1st and 2nd exhibition match questionnaire because none of the robots

damaged the playing surface in any noticeable way (i.e., this question did not discriminate between the robots). See Appendix D for the Competition questionnaire. See Appendix E for the evaluations provided by the participating teams prior to the Competition.

Each team participating (i.e., six teams participated and four teams did not) in this project completed a questionnaire for itself and each of the other participating teams by defining a range between 1 and 10 over which the robot in question could be expected to perform in each of the questions asked. Some exceptions to this are the following: team 2 had questionnaires filled out regarding its projected success but did not, itself, fill out questionnaires; team 6 filled out questionnaires for the 2nd exhibition and the competition but did not complete any questionnaires for the 1st exhibition.

3.2 Data tabulated from results of Exhibitions and Competition

Additionally, William Hardman, who observed each of the exhibitions and the competition, tabulated results. The results were tabulated by William Hardman's completion of one questionnaire for each of the participating teams during each of the three events. The difference (as compared to the questionnaires filled out by the teams prior to the three events) in this case was that the result for a specific question was assigned to only one bin, not a range of bins. The proceeding was accurate because these results were discrete physical outcomes, not projections of future success. This means that one bin in each question was assigned a value of unity (i.e., the one bin which best indicated the success/failure of the team to meet the performance requirement of that question).

Each question was weighted as previously discussed in Chapter 4. Each of the questions formed a part of a group (control, control and score, and score) of questions, with each group's total weighting summing to unity (as previously discussed). For a specific bin of a given question, the weighting of the question is multiplied by the probability of being in the bin. For each bin, this product is summed for each question in the group. The result is the probability of being in any bin for a given group. Each group was weighted with total weighting between the three groups summing to unity (as is previously discussed). Just as the questions within a group were reduced to corresponding group probabilities of being in specific bins, the groups were, then, reduced to an overall probability of a robot team being in a specific bin. These probabilities will be called results in all of the ensuing discussion. Results for Exhibitions 1 and 2 and the Competition are discussed in Chapter 5. Specifically, the results of Exhibition 1, Exhibition 2 and the Competition are shown in Figure 7, Figure 14, and Figure 20 respectively.

Chapter 4

How the data were weighted

4.1 Question weighting from Exhibitions 1 and 2

Weightings (i.e., a fraction of unity) were assigned to each of the questions within a group such that the sum of the weightings for a group is equal to unity. An exception is the control/score group with its two sub-goals (called categories 1 and 2), which each has weightings summing to unity. Lastly, each of the groups is assigned a weighting such that the sum of the group weightings is likewise equal to unity.

At the onset of the project, the weightings described above were assigned based upon expectations by the Maslab staff of how the robot teams would attempt to score and operate, in general. These expectations were most easily quantified in the score group where point values were assigned to each of the means by which points could be accrued in the matches. The point values assigned to various activities established a natural ranking or weighting to each of the score group questions (i.e., highest points correspond to highest weighting). Points were assigned as follows: placing a target in the home area (5 points), placing a target in a yellow score area (3 points), performing a waypoint signal meaning making a noise or mechanical signal upon arriving at a target location (1 point), arriving home after the end of the allotted 4 minutes (1 point), and moving after the allowed 4 minutes (-1 point).

After observing the first and second exhibition matches, it became clear that the contestants would attempt to take targets to the yellow score areas as opposed to the home area. Therefore, although more points were available for taking a target home, this scoring mechanism was weighted third in importance (weighting = 0.2) after placing targets in the yellow score areas (weighting = 0.4) and making a signal once arriving at a target or waypoint (weighting = 0.3). The last two scoring mechanisms for arrival at home in less than four minutes (weighting = 0.05) or moving after the allotted four minutes had expired (weighting = 0.05) were weighted only minimally. The contestants all showed themselves capable of not exceeding the allowed four minutes operation time; and, thus, none would be negatively affected by the loss of one point for exceeding the allowed time. Also, none of the competitors showed any indication that the goal was to end the matches with the robot having made its way back to home. Thus, this scoring mechanism was not chosen as a scoring means by any of the competitors and became non-discriminating and likewise was minimally weighted.

The control/score group (category 1) weightings initially were equal to 0.6 for arrival at targets, equal to 0.2 for detection of the targets, and equal to 0.2 for processing quickly in detecting targets and score areas. We assumed that having high reliability in detecting targets and short processing times for detecting targets were essentially different versions of the same question and thus merited the same weighting. The more quickly a robot processed the sensor data and determined a detection had occurred made for a much more timely signal and thus a more reliable robot. This was true as longer processing times meant more time between when a robot first obtained sensor data on a target and when it determined that a target had been sighted which also meant more of a

delay between first sensor data and time to signal the detection of the target. The robot's arrival at the target was the most easily measured (i.e., visually observed) means by which the control and scoring capability of the target can be measured; and, thus, was given the highest weighting.

The control/score group (category 2) weightings did not change throughout the project. The weightings were equal to 0.4 for getting to a yellow score area, equal to 0.3 for detaching the target from the robot into the yellow score area, equal to 0.15 for the fraction of the playing field explored, equal to 0.1 for reliability in detecting the score area, and 0.05 for attempting to grasp a target after failing to grasp it on the first attempt. The weightings most representative of success of the robot in this category were assigned to the robot's arrival at the yellow score area followed by the ability to detach the target in the score area. This preference was natural because the robot must first get to the score area else detaching the target became irrelevant. Percentage of playing field explored was not a direct link to success. However, the idea was that if more area were explored then the robot would be more successful; this was only true if the area was explored efficiently (i.e., long searches to identify targets, once near them, or random travel bypassing numerous targets are not desirable). Reliability in detecting the scoring areas was not quite as easily quantifiable a metric as actual arrival at the score areas and, though important, was just part of the process of arrival at a score area and must receive a lower weighting. The robots, with one exception, used the approach of dragging targets wedged in a bay beneath them, thus minimizing the importance of this question in discriminating between robots.

The control group weightings initially were equal to 0.4 for the 10 second required computing and movement delay prior to the start of the 4 minute play period (allowing for fair start for all robot competitors), equal to 0.3 for the robot's ability to avoid collisions with large objects (like walls in the match arena), equal to 0.2 for having a calibration time of less than 60 seconds prior to a match (i.e., time required to prepare the robot for the start of a match once placed in the home area) and equal to 0.1 for the ability to operate reliably using battery power. In the first and second exhibitions the robots all had difficulty especially with the ten second delay, collision with walls and the calibration time prior to the matches. At that stage of the project the robots automatically started at the push of their start buttons, collided frequently with walls and often stalled following a collision, and required significantly more calibration time than the 60 seconds allowed prior to a match. The weightings described above indicate the relative difficulty that the robots had in each performance area. Operating reliably using battery power appeared to be largely a non-discriminator between the robots and, thus, was weighted minimally.

4.2 Question weightings revised

Table 2: Initial and final values of the weightings for Control questions

Control (goal)	**Sub-goal weighting**	
	initial	final
Sub-goals		
calibration time	0.1875	0.025
10 sec required delay	0.3875	0.025
damages playing surface (penalty)	0.05	0.01
collision with objects in path	0.2875	0.8
reliable operation on battery power	0.0875	0.14
total	1	1

Table 3: Initial and final values of the weightings for Control and Scoring questions

Control and Scoring (goal) (two categories)		
	Sub-goal weighting	
Category 1 sub-goals	initial	final
arrival at targets	0.6	0.6
reliability in detecting targets	0.2	0.1
long processing time in detecting targets and score areas	0.2	0.3
total	1	1
	Sub-goal weighting	
Category 2 sub-goals	initial	final
arriving at yellow score areas	0.4	0.4
detaching target in score areas	0.3	0.3
attempting to grasp target again if failed on first attempt	0.05	0.05
reliability in detecting score areas	0.1	0.1
percentage of playing area explored	0.15	0.15
total	1	1

Table 4: Initial and final values of the weightings for Scoring questions

Scoring (goal)		
	Sub-goal weighting	
Sub-goals	initial	final
robot performs signal at waypoint (target)	0.3125	0.16
robot places target in score area	0.4125	0.11
robot places target in home area	0.2125	0.66
robot returns home in < 4 min. allowed	0.0625	0.07
total	1	1

After observing the final match, the competition, it became apparent that the weightings had to be revised significantly in the control and score groups, as shown in Table 2 and Table 4, respectively, with only minor revisions being made in the control/score group (category 1) shown in Table 3. These revisions were made in order to represent more accurately the importance of those performance attributes (or goals), which reflected best the results of the competition. By the time of the competition, the robots were in the best position to exhibit all of the skills developed over the course of the project as opposed to only showcasing the results of the progressive design enhancements following the first and second exhibitions. Design and programming choices that the winning competitors had made reflected different priorities among the set of performance goals, and, thus, dictated different weights in forecasting accurately the outcomes of the competition. This meant that weights were adopted such that observation of the attributes contributing most to the competition results combined with the actual competition results (i.e., ranking of competitors at the conclusion of the competition) could provide a picture closely resembling the competition outcome.

The most significant weighting revisions occurred in the score group shown in Table 4. The robots placing highest in the competition chose to take targets home as opposed to

the yellow score areas, as was done in the first and second exhibitions. This was the major mechanism used by competitors to score highly in the competition. Thus, the weighting was changed from 0.2 to 0.65. The second major means used to score was the making of a signal upon arrival at a target, dictating a weighting change from 0.4 to 0.15. The third most important scoring method was that of taking targets to the scoring area. However, this was largely unused, and, thus, was weighted accordingly at 0.1 (previously at 0.3). The lowest weightings were assigned to arrival at home in less than four minutes (going to 0.07 from 0.05) which was not a competitor priority but which was slightly more important than moving after the four minutes allowed (0.03). None of the robots moved after the allowed four minutes. Thus, the question of performance in this area became a non-discriminator and, therefore, was weighted negligibly.

The control group weighting revisions largely made all but one question in this group irrelevant as discriminators. Collision with large objects (i.e., walls) became the telling question within this group because robots either wasted significant time or stalled completely once a collision had occurred. Additionally, robots stalled occasionally without having collided with anything. These stalls were accounted for here also. The weighting for collisions and stalls was assigned a value of 0.8. (0.3 previously). Reliable operation using battery power, though largely non-discriminating, was discriminated enough in the competition to merit a weighting of 0.14 (previously 0.1). The questions regarding the calibration time requirement and the ten second delay prior to robot movement and computation proved to be non-discriminators, and were both weighted negligibly at 0.03 (previously 0.2 and 0.4 respectively).

The control/score group (category 1) weightings were only minimally changed. The arrival at the targets maintained its precedence with a weighting of 0.6, but the processing time needed to detect targets and score areas was raised from 0.2 to 0.3 to account for the fact that this question included detection of not only targets but also score areas (a fact neglected in the first weighting assessment). Similarly, the weighting for reliability in detection of targets was lowered from 0.2 to 0.1 in order to account for the incorrect weighting assessment just described.

4.3 Group weightings

Table 5: Initial and final values of the weightings for Control and Scoring sub-goals

Control and Scoring (goal)			
	weightings		
		initial	final
category 1 sub-goals		0.5	0.3
category 2 sub-goals		0.5	0.7

Table 6: Initial and final values of the weightings for Control, Control and Scoring, and Scoring goals

Goal weightings		
	initial	final
Control (goal)	0.3	0.1
Control and Scoring (goal)	0.2	0.3
Scoring (goal)	0.5	0.6
total	1	1

The groups were initially assigned weightings of 0.3 for control, 0.2 for control and score (with 0.5 assigned to each of categories 1 and 2 as shown in Table 5), and 0.5 for score. In order to accurately represent the results of the competition the group weightings were changed to 0.1 for control, 0.3 for control and score (with 0.3 for category 1 and 0.7 for category 2 as shown in Table 5), and 0.6 for score. See Table 6 for the initial and final group weightings.

In the control group only one question appeared to be a real discriminator, as discussed previously. Furthermore, this question was not a real descriminator in points scored by the victorious robots; hence a weight reduction of 0.2 was merited in this group (from 0.3 to 0.1).

Based on how points were scored by the robots in the competition, no different (from that in the first and second exhibitions) discrimination appeared to exist for the control and score group compared to the score group. Therefore, the difference in weightings, as previously assigned, was maintained. This meant that the 0.2 weight reduction for the control group was assigned equally between the control and score group and the score group (see Table 6).

The control and score category 1 and 2 weightings (see Table 5) should not have been equal given the fact that more questions are asked in category 2 than in category 1 (5 versus 1) as well as the fact that questions were asked which discriminated more effectively the various levels of success displayed by the robots (compared to category 1 questions).

Chapter 5

Results of the various robot matches

Throughout Chapter 5, the overall ranking will specify the ranking of the teams as evaluated by a specific team or by the results of a match (i.e., exhibitions 1 or 2 or the competition). The overall ranking in all of these cases is taken from the expected value of the bin number for each team. The bin number in which a team is expected to be found is calculated using the equation below for expected value, E(y), given a discrete number of bins (i.e., n = 10).

Equation 1: Expected value (i.e., number of the bin in which a team is expected to be found) immediately follows and then is defined below.

Equation 1: Expected value [4]

$$E(y) := \sum_{i=1}^{n} y_i \cdot p_i$$

The probability, p_i, for any team, of being in each of the ten bins, respectively, (n = 10) has been assigned. Multiplying the probability, p_i, of being in a bin by the bin number, y_i, and then summing over the ten products gives the expected bin in which a team should be found. The higher the bin number, the better the team is expected to perform, and the higher ranking that is achieved.

Also, throughout Chapter 5, the word ranking means a ranking of highest to lowest probabilities of a team being in a specific bin. Therefore, this ranking is also a ranking (1^{st} to last) of teams' performance for that specific bin.

Table 7: Far left column indicates the team whose evaluations produced the Exhibition 1, Exhibition 2 and Competition rankings as a function of expected bin in which a team should be found. The teams are ranked highest to lowest with the highest at the top and lowest at the bottom. To the right of the Exhibition 1, Exhibition 2 and Competition rankings are the results, in the form of team rankings also, for each of the respective events.

		Rankings for	Results	Rankings for	Results	Rankings for	Results
		Exhibition 1	Exhibition 1	Exhibition 2	Exhibition 2	Competition	Competition
Team 1	rank	team		team	team	team	team
	1	2	assumed average	1	10	5	2
	2	5	performance for	3	3	10	10
	3	10	all teams	6	2	3	1
	4	3		10	6	6	6
	5	6		5	1,5	2	3
	6	1		2		1	5
Team 3	rank	team		team	team	team	team
	1	10	assumed average	10	10	3	2
	2	6	performance for	2	3	10	10
	3	2	all teams	6	2	2	1
	4	1		1	6	6	6
	5	3		5	1,5	1	3
	6	5		3		5	5
Team 5	rank	team		team	team	team	team
	1	3	assumed average	3	10	10	2
	2	2	performance for	2	3	1	10
	3	5	all teams	1	2	3	1
	4	6		10	6	6	6
	5	10		6	1,5	2	3
	6	1		5		5	5
Team 6	rank ****	team		team	team	team	team
	1	5	assumed average	5	10	10	2
	2	1	performance for	1	3	1	10
	3	3	all teams	3	2	3	1
	4	2		2	6	6	6
	5	6		6	1,5	2	3
	6	10		10		5	5
****	Team 6 eval	uations used	from Exhibition 2	since none	provided from	Exhibition 1.	
Team 10	rank	team		team	team	team	team
	1	6	assumed average	2	10	10	2
	2	2	performance for	10	3	2	10
	3	3	all teams	3	2	3	1
	4	10		5	6	6	6
	5	5		1	1,5	5	3
	6	1		6		1	5

5.1 Exhibition 1

5.1.1 Exhibition 1 projections (of success) made by participating teams

In Figure 2 (on the next page) see the evaluation by team 1 of the probability of success for the teams in exhibition 1. Team 1 predicted that the overall ranking, from first to last, would be team 2, team 5, team 10, team 3, team 6, and finally team 1 (see Table 7). Given the minor variation in probabilities for all teams in either bin 1 or bin 10, it is apparent that team 1 believed that all teams had an approximately equal likelihood of being in bin 1. Likewise, team 1 believed all teams had an approximately equal likelihood of being in bin 10.

Team 1 predicted the ranking (i.e., first to last) for bin number 10 would be team 2, team 3, team 5, team 6, team 10, and finally team 1. For bin number 5, team 1 predicted the ranking (again first to last) as team 10, team 1, team 6, team 3, team 5, and finally team 2. In bin number 1, team 1 predicted the ranking (also first to last) as team 6, team 1, team 5, team 10, team 2, and lastly team 3.

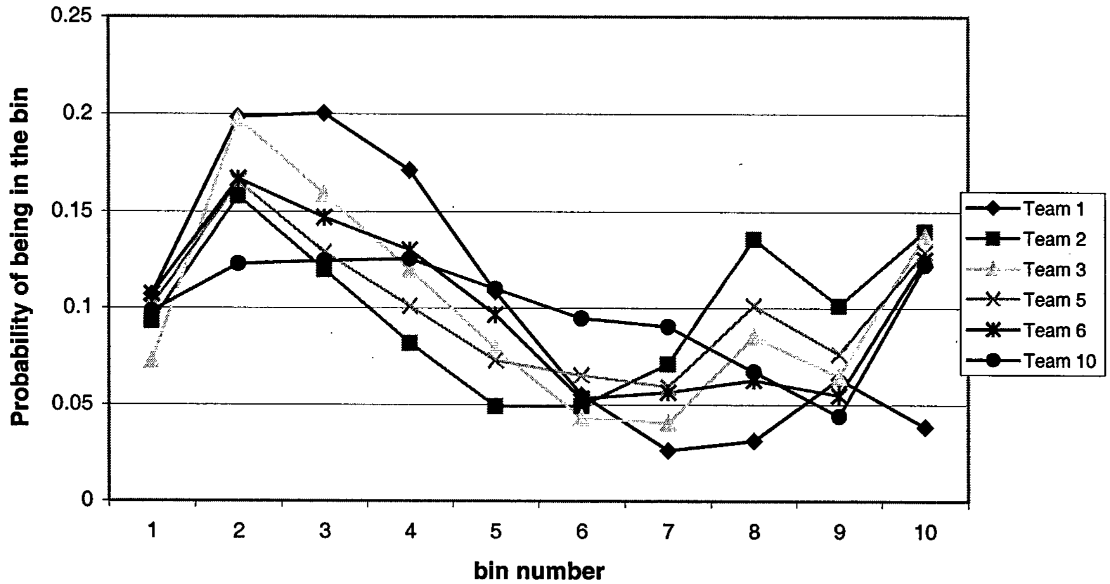

Figure 2: Prior evaluation by team 1 of the probability of success for the teams in exhibition 1.

In Figure 3 below see the evaluation by team 3 of the probability of success for the teams in exhibition 1. Team 3 predicted that the overall ranking, from first to last, would be team 10, team 6, team 2, team 1, team 3, and finally team 5(see Table 7). . Figure 3 shows that teams 10, 6, 2, and 1 were expected to do well but that teams 3 and 5 were expected to do poorly.

Team 3 predicted the ranking (i.e., first to last) for bin number 10 would be team 2, team 10, team 6, team 1, team 3, and finally team 5. For bin number 5, team 3 predicted the ranking (again first to last) as team 1, team 2, team 10, team 6, team 5, and finally team 3. Lastly, in bin number 1, team 3 predicted the ranking (also first to last) as team 5, team 3, team 1, and teams 2, 6, and 10.

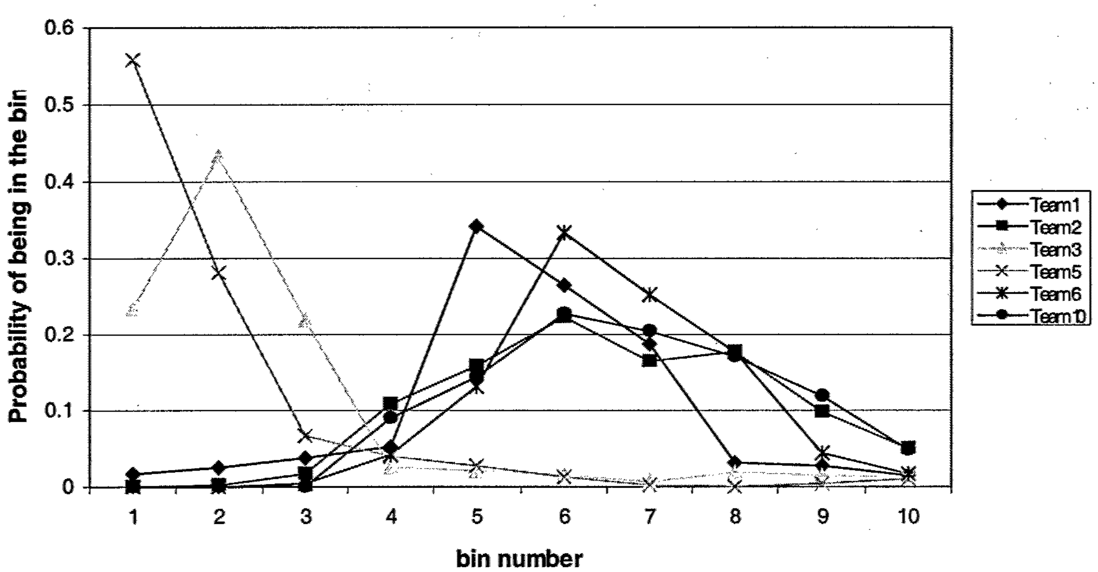

Figure 3: Prior evaluation by team 3 of the probability of success for the teams in exhibition 1.

In Figure 4 below see the evaluation by team 5 of the probability of success for the teams in exhibition 1. Team 5 predicted that the overall ranking, from first to last, would be team 3, team 2, team 5, team 6, team 10, and finally team 1(see Table 7). . As shown in Figure 4, Team 5 expected largely poor performance from all but team 3.

Team 5 predicted the ranking (i.e., first to last) for bin number 10 would be team 2, team 5, team 1, team 3, team 10, and finally team 6. For bin number 5, team 5 predicted the ranking (again first to last) as team 3, team 10, teams 2 and 5, team 6, and finally team 1. In bin number 1, team 5 predicted the ranking (also first to last) as team 1, team 6, teams 5 and 10, team 2, and lastly team 3.

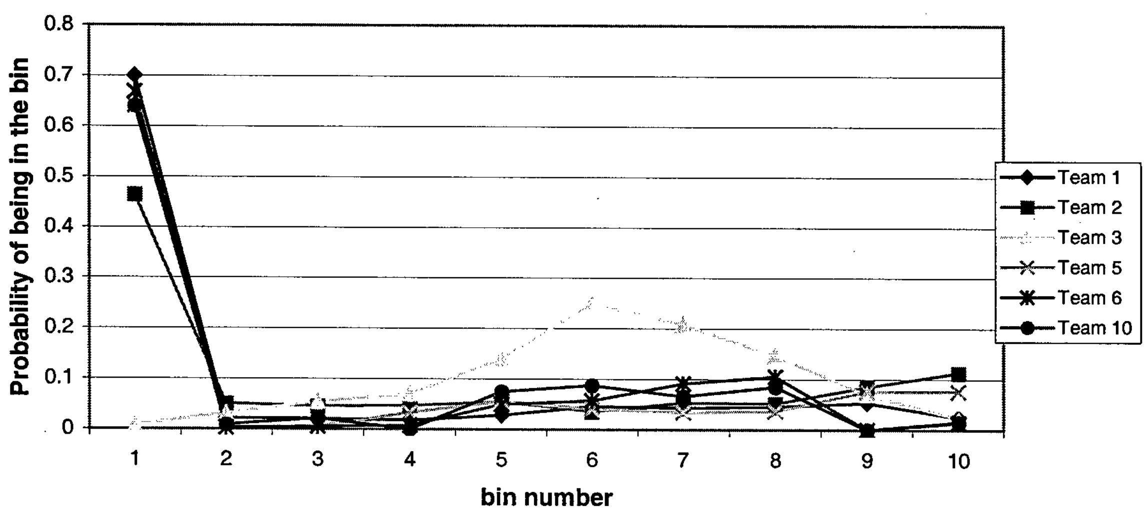

Figure 4: Prior evaluation by team 5 of the probability of success for the teams in exhibition 1.

Team 6 provided no evaluations for Exhibition 1. Instead of disregarding Team 6 evaluations for Exhibition 2 and the Competition, the Team 6 evaluations for Exhibition

2 were substituted for the, non-existent, evaluations for Exhibition 1. This fact minimizes the value of team 6 evaluations.

In Figure 5 below see the evaluation by team 10 of the probability of success for the teams in exhibition 1. Team 10 predicted that the overall ranking, from first to last, would be team 6, team 2, team 3, team 10, team 5, and finally team 1(see Table 7). . As shown in Figure 5, team 10 expected all teams to have moderate success as shown by the approximately equal probabilities (0.57 to 0.6) of being in bin 5.

Team 10 predicted the ranking (i.e., first to last) for bin number 10 would be team 3, team 6, team 2, team 10, and finally teams 1 and 5. For bin number 5, team 10 predicted the ranking (again first to last) as team 10, teams 3 and 5, team 2, team 1, and finally team 6. In bin number 1, team 10 predicted the ranking (also first to last) as team 10, team 5, team 1, and lastly teams 2,3 and 6.

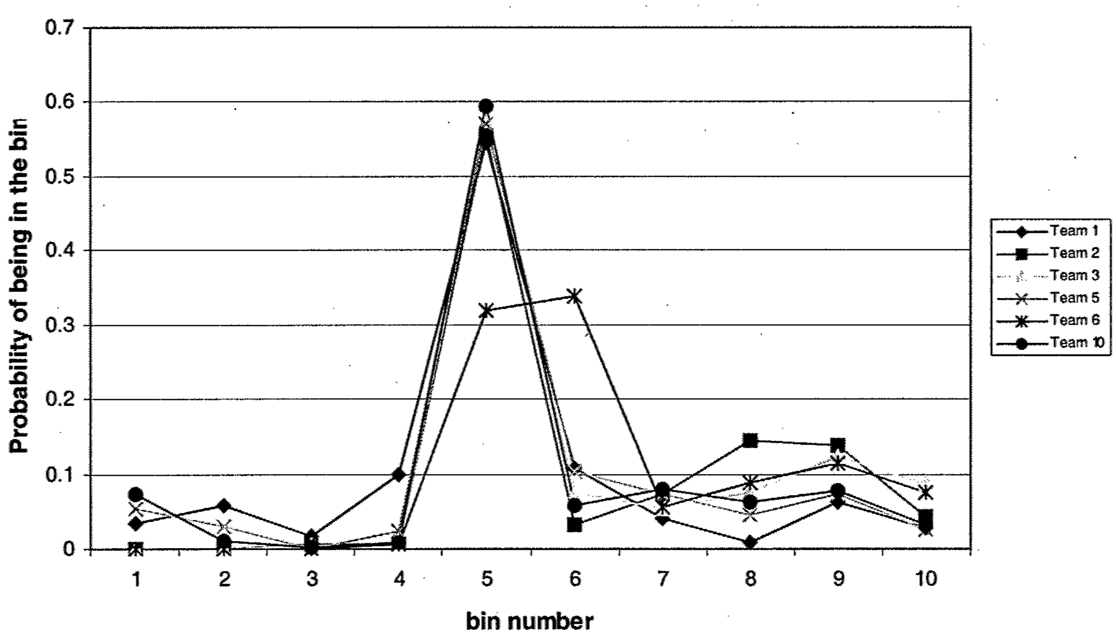

Figure 5: Prior evaluation by team 10 of the probability of success for the teams in exhibition 1.

5.1.2 Results of exhibition 1

The only importance assigned to Figure 6 below is the fact that none of the participating teams was prepared to have its robots take part in exhibition 1. This means that scores of one (i.e., the lowest score) were officially assigned to each of the participating teams. Each of the participating teams was, however very close to being prepared for exhibition 1, a fact that is useful in the description for Figure 7. Note that although it appears in Figure 6 that only team 2 participated in exhibition 1 the one line of the plot applies to all 6 participating teams.

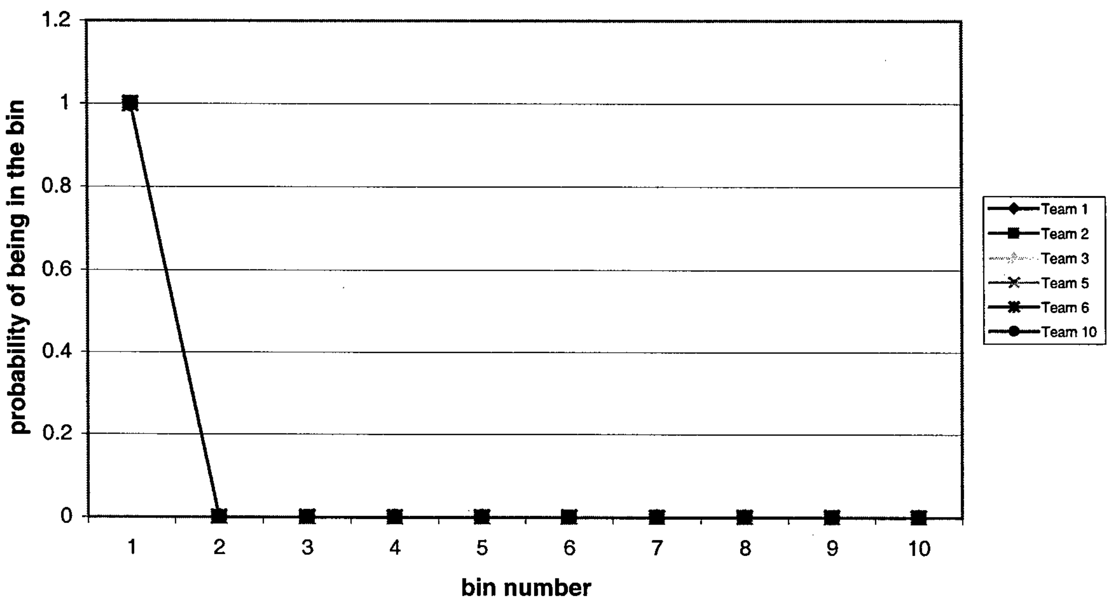

Figure 6: Exhibition 1 results (participating teams). The one line of the plot applies to all 6 participating teams and merely indicates that the participating teams were not prepared for Exhibition 1.

Note in Figure 7 below the results for exhibition 1 from the teams not participating in our project are plotted. This is done because none of the participating teams have scores that are useful for the Bayesian approach to projecting the success of the robots in successive events, which is discussed in Chapter 6. Therefore, an initial result for all of the participating teams was taken as the average (in each bin) of the non-participating team results such that a useful starting point for further projections could be had. These bin-wise averages were taken to be the evidence obtained from tests of the likely performance of a typical team that was able to get its robot working. Use of this evidence permits Bayesian projections of team-specific performance for exhibition 2 and for the competition evaluations to be closer to reality than if scores of unity in bin # 1, as shown in Figure 6, were assigned to each of the participating teams for exhibition 1.

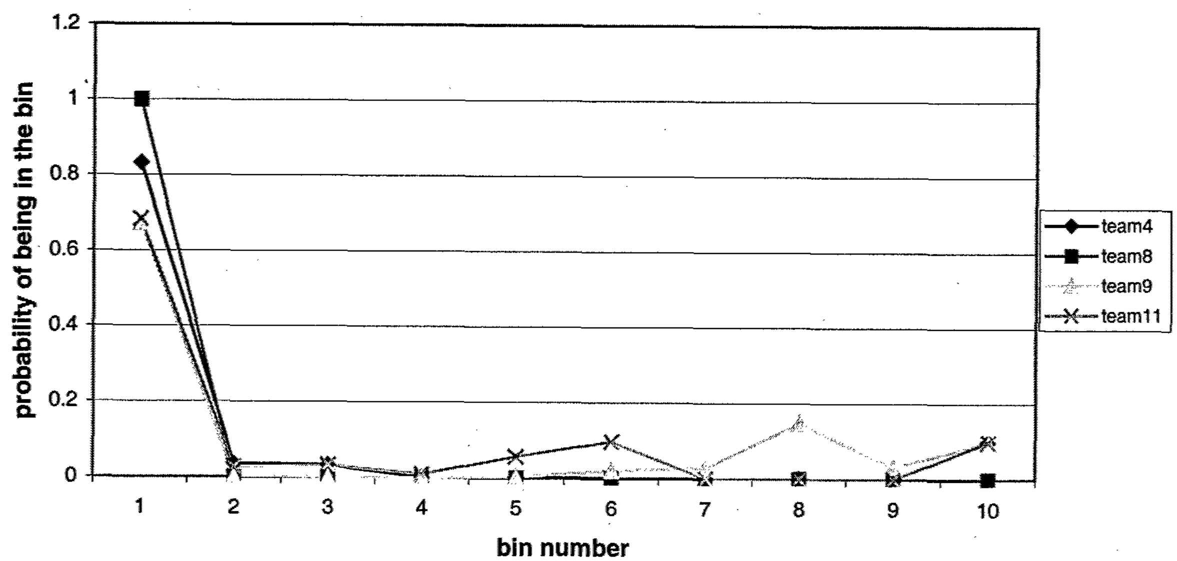

Figure 7: Exhibition 1 results (non-participating teams)

Figure 8 below graphically illustrates the assumed values of the results for each of the participating teams. These values (bin-wise averages of the non-participating teams results for Exhibition 1, see Figure 7) were assumed since the participating teams were not prepared for Exhibition 1.

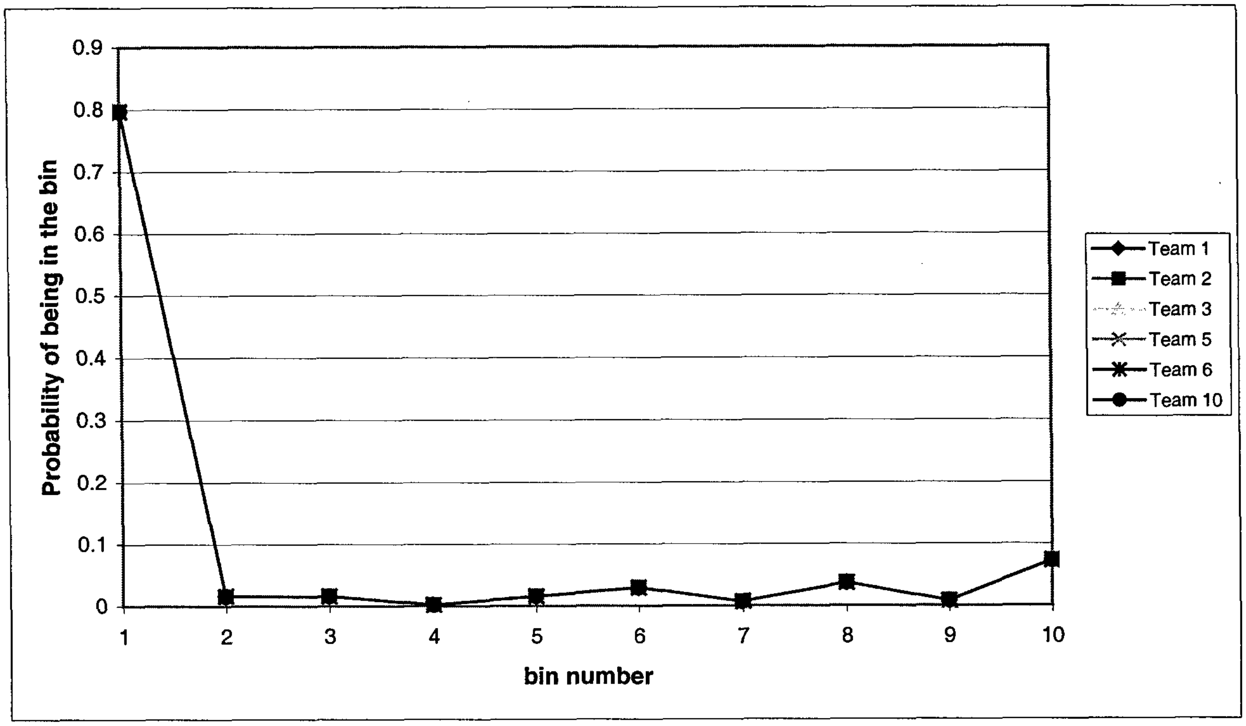

Figure 8: Assumed Exhibition 1 results for each of the participating teams. These results are the bin-wise averages of the data presented in Figure 7.

5.1.3 A general description of faults of the robots in Exhibition 1 follows.

Manual starts of robots were required. A manual start means, simply, that a button is pushed and the robot begins computing and/or moving immediately after the start button is depressed (i.e., 10-second delay not met as discussed below).

Most robots began moving or computing prior to the required 10-second delay at the beginning of the event. The 10-second delay is meant to ensure that the robots would begin the event at exactly the same time, demonstrating control of the robot.

The robots spent considerable time scanning for targets. Once a target was approached many robots were still unable to signal by a noise or mechanical action that the robot is within four inches of the target.

Most sensors (e.g., ultrasound or infrared range finders) were not operational. For most robots vision was the only sensor as of yet at their disposal. This made it possible to detect the targets but impossible to stop (because only the code for target detection and movement toward the detected target, not what to do once target was approached, was written at this point) and signal; hence robots just drove through the targets without stopping. Once the robot drove over the target it lost sight of the target and began looking for other targets.

5.2 Exhibition 2

5.2.1 Exhibition 2 projections (of success) made by participating teams

In Figure 9 (next page) see the evaluation by team 1 of the probability of success for the teams in exhibition 2. Team 1 predicted that the overall ranking, from first to last,

would be team 1, team 3, team 6, team 10, team 5, and finally team 2 (see Table 7). This prediction was accurate for two teams, teams 3 and 5. Teams 5 and 2 were expected to perform poorly; teams 6 and 10 were expected to be about average performers; while teams 1 and 3 were expected to be good performers.

Team 1 predicted the ranking (i.e., first to last) for bin number 10 would be team 1, team 3, team 10, team 5, team 6, and finally team 2. For bin number 5, team 1 predicted the ranking (again first to last) as team 2, team 6, team 1, team 10, team 3, and finally team 5. In bin number 1, team 1 predicted the ranking (also first to last) as team 10, team 3, team 5, team 6, team 2, and lastly team 1.

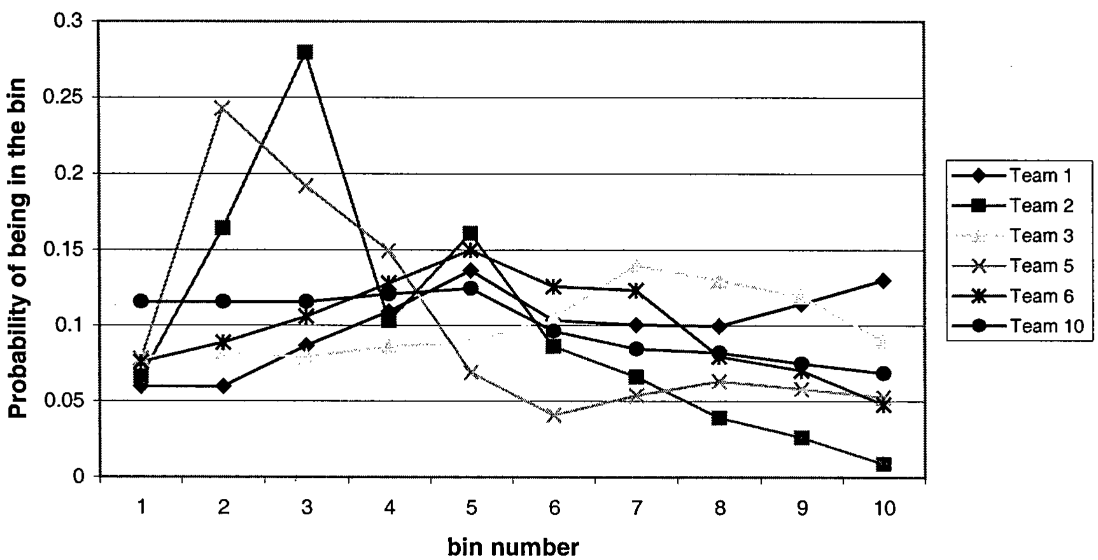

Figure 9: Evaluation by team 1 of the probability of success for the teams in exhibition 2.

In Figure 10 below see the evaluation by team 3 of the probability of success for the teams in exhibition 2. Team 3 predicted that the overall ranking, from first to last, would be team 10, team 2, team 6, team 1, team 5, and finally team 3(see Table 7). Team 3 was accurate on two teams (teams 10 and 5). Team 3 expects poor performance from team 3; average performance from teams 1, 5, and 6; and good performance from teams 2 and 10.

Team 3 predicted the ranking (i.e., first to last) for bin number 10 would be team 10, team 2, teams 1,5 and 6, and finally team 3. For bin number 5, team 3 predicted the ranking (again first to last) as team 6, team 5, team 1, team 3, team 2, and finally team 10. In bin number 1, team 3 predicted the ranking (also first to last) as team 3, team 1, and teams 2, 5, 6, and 10.

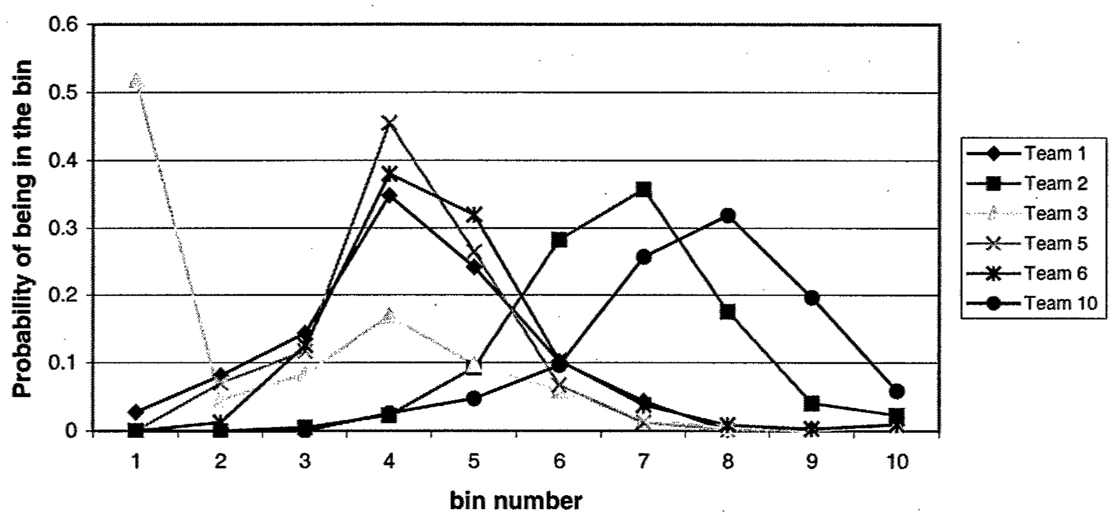

Figure 10: Evaluation by team 3 of the probability of success for the teams in exhibition 2.

In Figure 11 below see the evaluation by team 5 of the probability of success for the teams in exhibition 2. Team 5 predicted that the overall ranking, from first to last, would be team 3, team 2, team 1, team 10, team 6, and finally team 5(see Table 7). This prediction was accurate only for team 5. Team 5 predicted poor performance from team 5; average performance from teams 1, 2, 6, and 10; and good performance from team 3.

Team 5 predicted the ranking (i.e., first to last) for bin number 10 would be team 3, team 5, team 2, team 1, team 6, and finally team 10. For bin number 5, team 5 predicted the ranking (again first to last) as team 1, team 2, team 10, team 6, team 3, and finally team 5. In bin number 1, team 5 predicted the ranking (also first to last) as team 5, team 6, team 10, and teams 1,2, and 3.

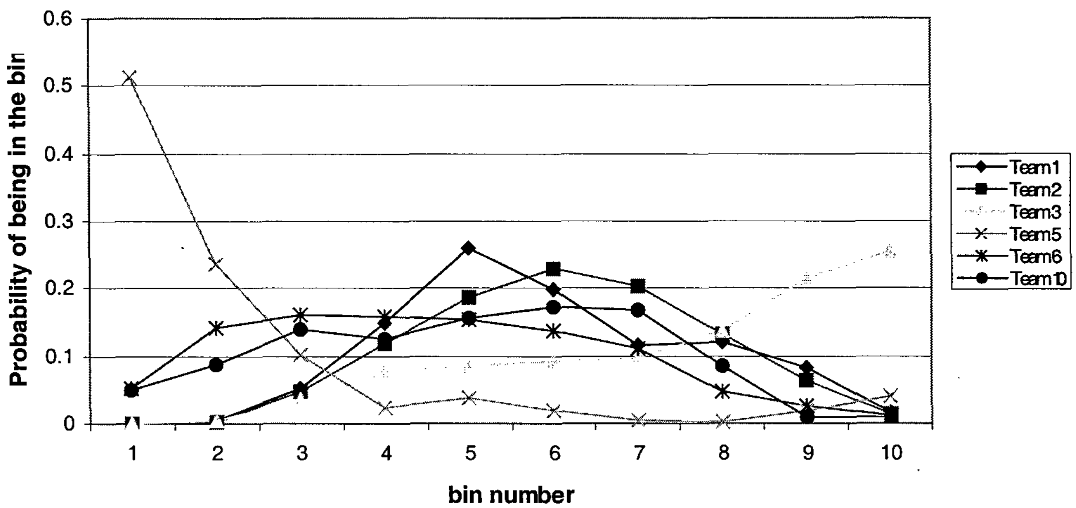

Figure 11: Evaluation by team 5 of the probability of success for the teams in exhibition 2.

In Figure 12 below see the evaluation by team 6 of the probability of success for the teams in exhibition 2. Team 6 predicted that the overall ranking, from first to last, would be team 5, team 1, team 3, team 2, team 6, and finally team 10 (see Table 7). Team 6 was accurate on none of these predictions. Team 6 predictions were somewhat inconsistent: team 10 poor or good; team 5 poor, average or good; team 3 poor or good; teams 2 and 6 about average, team 1 good.

Team 6 predicted the ranking (i.e., first to last) for bin number 10 would be team 5, team 10, team 1, team 3, team 6, and finally team 2. For bin number 5, team 6 predicted the ranking (again first to last) as team 1, team 6, team 2, team 5, team 3, and finally team 10. In bin number 1, team 6 predicted the ranking (also first to last) as team 10, team 5, team 3, team 2, team 6, and lastly team 1.

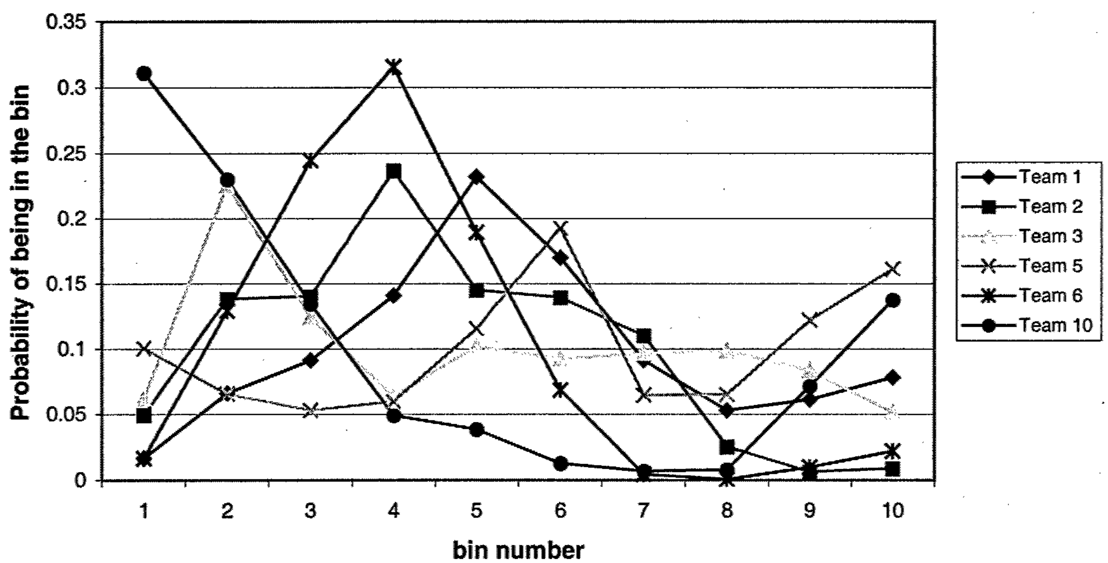

Figure 12: Evaluation by team 6 of the probability of success for the teams in exhibition 2.

In Figure 13 below see the evaluation by team 10 of the probability of success for the teams in exhibition 2. Team 10 predicted that the overall ranking, from first to last, would be team 2, team 10, team 3, team 5, team 1, and finally team 6 (see Table 7). Team 10 was accurate on one team (team 1). Team 10 predicted poor performance from teams 1, 3, 5, 6 and 10 and good performance from team 2.

Team 10 predicted the ranking (i.e., first to last) for bin number 10 would be team 10, team 3, team 5, team 2, team 1, and finally team 6. For bin number 5, team 10 predicted the ranking (again first to last) as team 2, team 6, team 5, team 1, team 10, and finally team 3. In bin number 1, team 10 predicted the ranking (also first to last) as team 1, team 6, team 3, team 5, team 10, and lastly team 2.

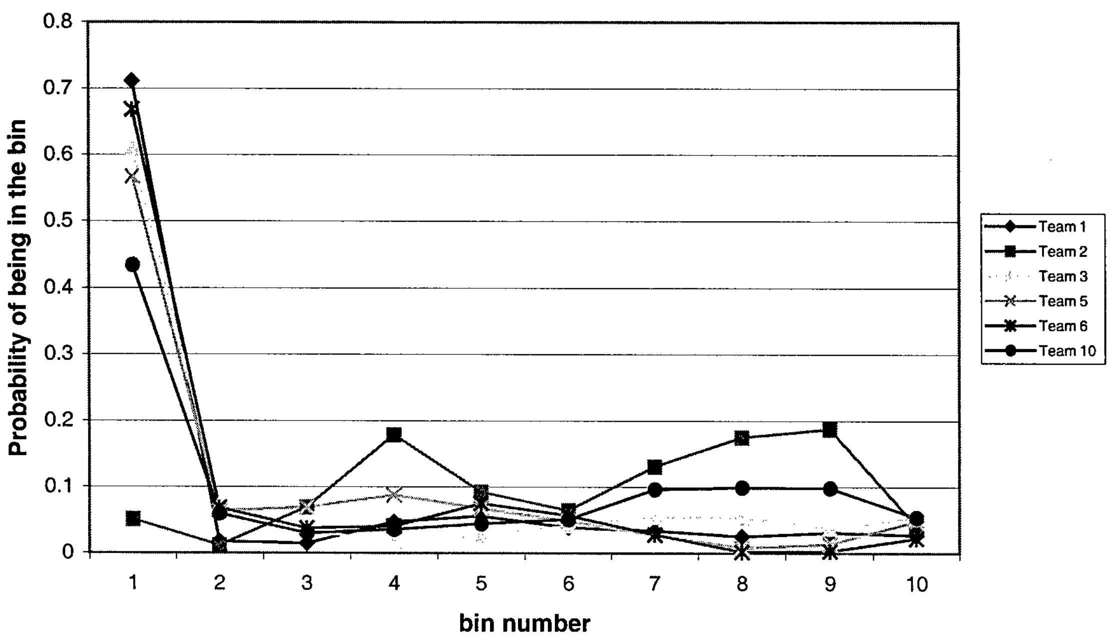

Figure 13: Evaluation by team 10 of the probability of success for the teams in exhibition 2.

5.2.2 Results of exhibition 2

In Figure 14 below see the results for the teams in exhibition 2. The overall ranking, from first to last, was team 10, team 3, team 2, team 6, and teams 1 and 5 (see Table 7). The results show poor performance for teams 1, 2, 6, and 5; slightly better than poor performance for team 3; and still slightly better performance from team 10. Summarizing, poor results were seen from all teams.

The results ranking (i.e., first to last) for bin number 10 was teams 2 and 10, teams 3 and 6, and teams 1 and 5. For bin number 5, the ranking (again first to last) was team 3, team 2, team 6, and teams 1, 5, and 10. In bin number 1, the ranking (also first to last) was teams 1 and 5, team 6, team 2, team 3, and lastly team 10.

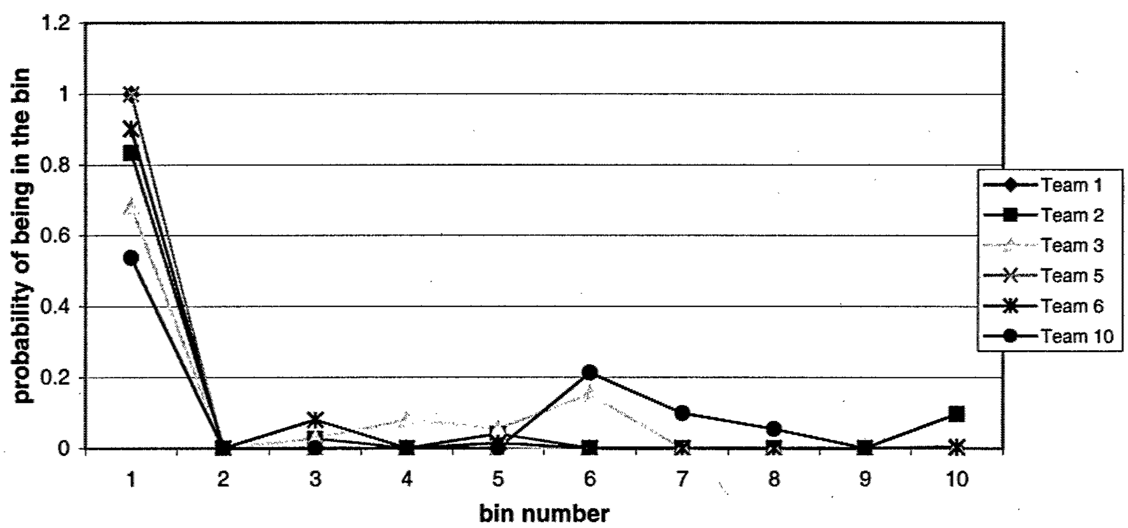

Figure 14: Exhibition 2 results (participating teams)

5.2.3 A general description of faults of the robots in Exhibition 2 follows. The robots in Exhibition 2 were still ramping up to attaining the final capabilities that will be exhibited in the Competition.

The robots were not moving for a long time after start of the four-minute contest period. Worse, the robots often moved only minimally from the home position (i.e., starting point) and then stalled requiring a restart of the robot's computer.

The robots still spent considerable time scanning for targets. Upon arrival at a target, some robots were still not signaling. Additionally, some robots did not pick up the targets but, rather, spun them endlessly instead. What targets were picked up were not readily detached in scoring areas.

5.3 The competition

5.3.1 Competition projections (of success) made by participating teams

In Figure 15 below see the evaluation by team 1 of the probability of success for the teams in the competition. Team 1 predicted that the overall ranking, from first to last, would be team 5, team 10, team 3, team 6, team 2, and finally team 1 (see Table 7). Team 1 was accurate on two teams (teams 6 and 10). Team 1 predicted poor performance from teams 1 and 2 and good performance from teams 1, 5, 6 and 10.

Team 1 predicted the ranking (i.e., first to last) for bin number 10 would be team 10, team 3, team 6, team 1, team 5, and finally team 2. For bin number 5, team 1 predicted the ranking (again first to last) as team 2, team 5, team 3, team 6, team 10, and finally team 1. In bin number 1, team 1 predicted the ranking (also first to last) as team 1, team 2, team 10, team 6, team 3, and lastly team 5.

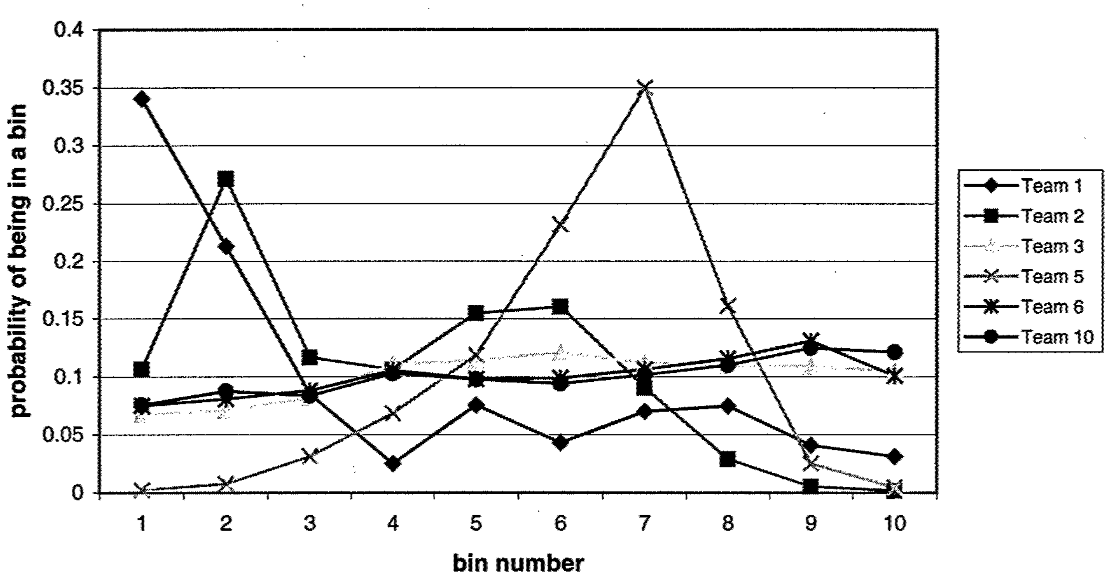

Figure 15: Evaluation by team 1 of the probability of success for teams in the competition.

In Figure 16 below see the evaluation by team 3 of the probability of success for the teams in the competition. Team 3 predicted that the overall ranking, from first to last, would be team 3, team 10, team 2, team 6, team 1, and finally team 5 (see Table 7). Team 3 was accurate on two teams (6 and 10). Team 3 expected poor performance from team 5; below average performance from teams 1, 2, and 6; average performance from team 10; and good performance from team 3.

Team 3 predicted the ranking (i.e., first to last) for bin number 10 would be team 3, team 10, team 2, teams 1 and 5, and finally team 6. For bin number 5, team 3 predicted the ranking (again first to last) as team 6, team 1, team 2, team 3, team 10, and finally team 5. In bin number 1, team 3 predicted the ranking (also first to last) as team 5, team 2, and teams 1, 3, 6, and 10.

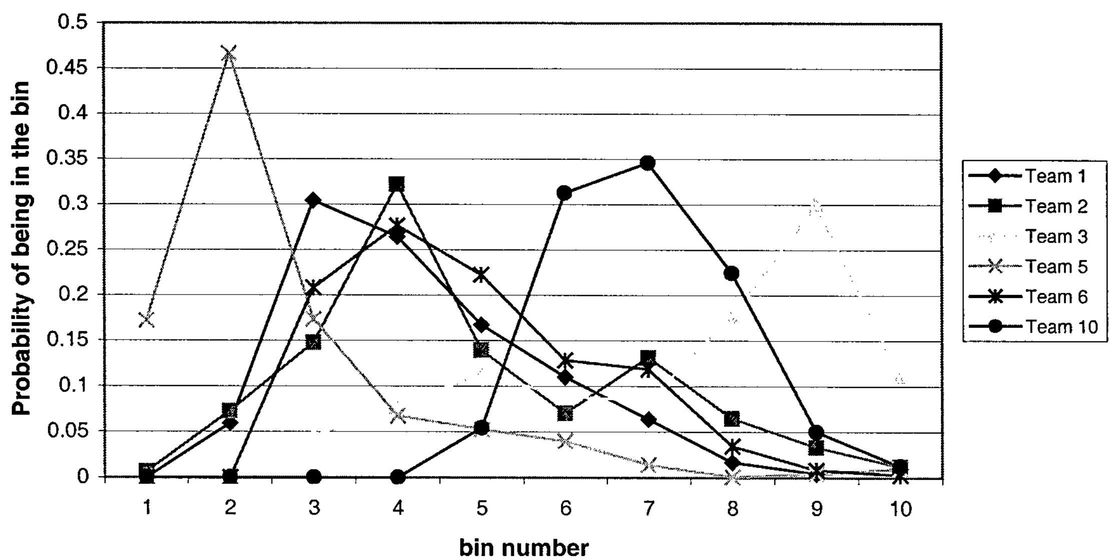

Figure 16: Evaluation by team 3 of the probability of success for teams in the competition.

In Figure 17 below see the evaluation by team 5 of the probability of success for the teams in the competition. Team 5 predicted that the overall ranking, from first to last, would be team 10, team 1, team 3, team 6, team 2, and finally team 5 (see Table 7). Team 5 was accurate on one team (team 6). Team 5 predictions were somewhat inconsistent: team 1 poor or good, team 6 poor or good, team 2 poor, team 5 poor, team 3 average, and team 10 good.

Team 5 predicted the ranking (i.e., first to last) for bin number 10 would be team 10, team 6, team 1, team 3, team 2, and finally team 5. For bin number 5, team 5 predicted the ranking (again first to last) as team 3, team 6, team 5, team 2, team 1, and finally team 10. In bin number 1, team 5 predicted the ranking (also first to last) as team 6, team 2, team 5, team 1, team 3, and lastly team 10.

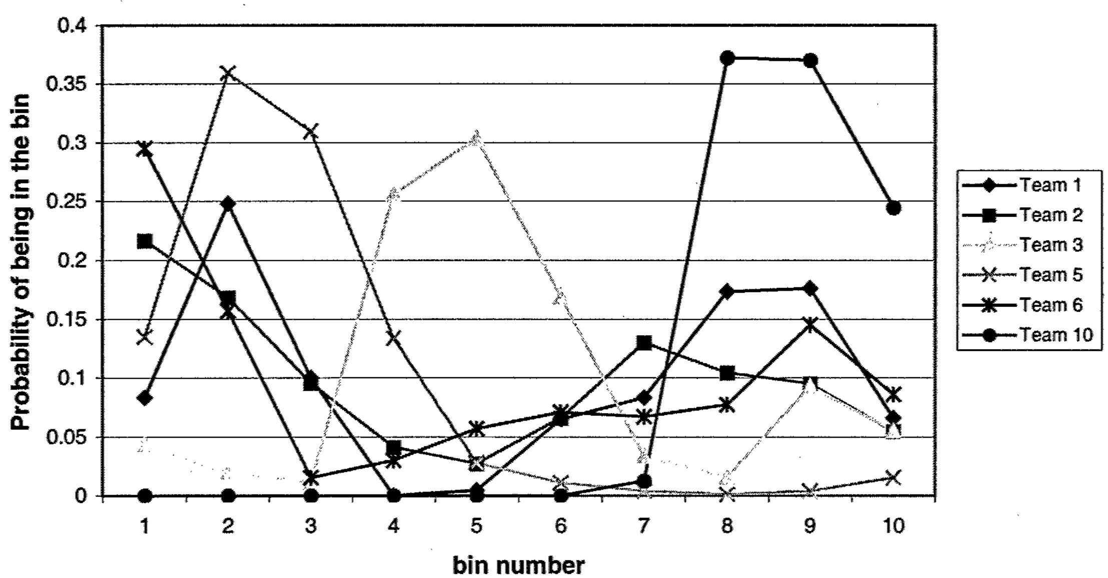

Figure 17: Evaluation by team 5 of the probability of success for teams in the competition.

In Figure 18 below see the evaluation by team 6 of the probability of success for the teams in the competition. Team 6 predicted that the overall ranking, from first to last, would be team 10, team 1, team 3, team 6, team 2, and finally team 5 (see Table 7). Team 6 was accurate for two teams (teams 5 and 6). Team 6 predictions were somewhat inconsistent: team 2 poor or good, team 6 poor or good, team 10 poor or good, team 5 poor, team 3 average, and team 1 average.

Team 6 predicted the ranking (i.e., first to last) for bin number 10 would be team 6, team 10, team 2, team 3, team 1, and finally team 5. For bin number 5, team 6 predicted the ranking (again first to last) as team 1, team 6, team 5, team 3, team 2, and finally team 10. In bin number 1, team 6 predicted the ranking (also first to last) as team 6, team 10, team 5, team 2, team 3, and lastly team 1.

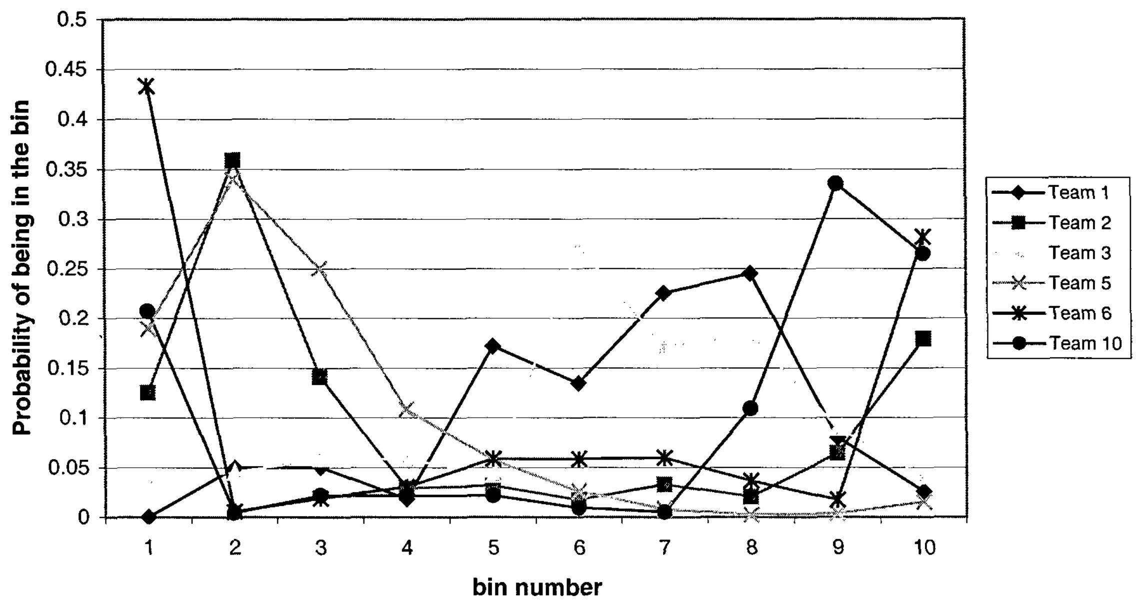

Figure 18: Evaluation by team 6 of the probability of success for teams in the competition.

In Figure 19 below see the evaluation by team 10 of the probability of success for the teams in the competition. Team 10 predicted that the overall ranking, from first to last, would be team 10, team 2, team 3, team 6, team 5, and finally team 1 (see Table 7). Team 10 was accurate on one team (team 6). Team 10 predicted poor performance for team 1, poor performance for team 5, average performance for team 3, average performance for team 6, good performance for team 10, and good performance by team 2.

Team 10 predicted the ranking (i.e., first to last) for bin number 10 would be team 10, team 2, team 3, and teams 1, 5, and 6. For bin number 5, team 10 predicted the ranking (again first to last) as team 6, team 3, team 1, team 2, team 5, and finally team 10. In bin number 1, team 10 predicted the ranking (also first to last) as team 1, team 10, team 5, team 2, team 3, and lastly team 6.

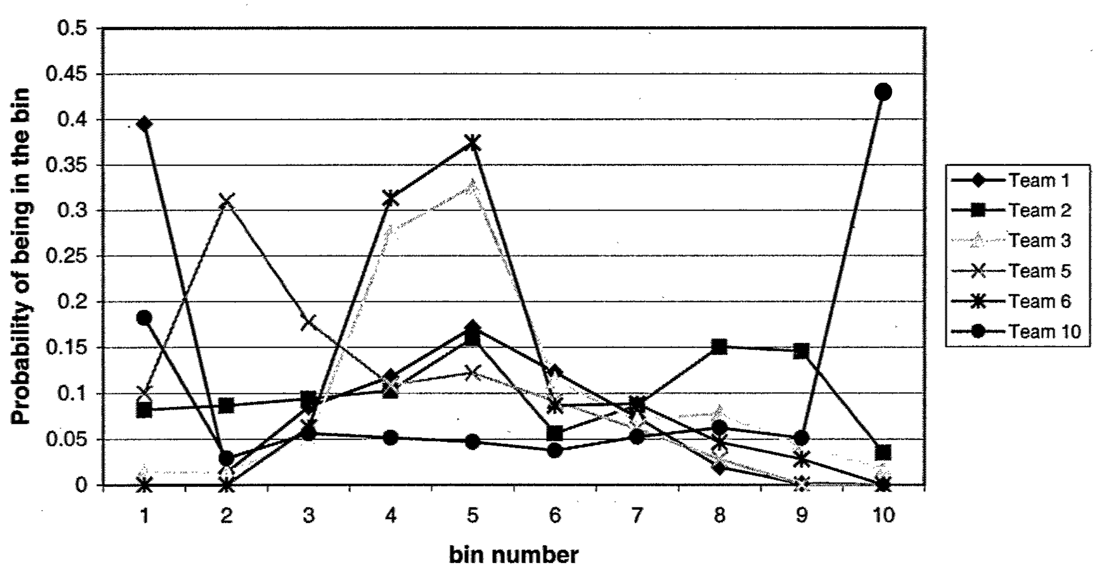

Figure 19: Evaluation by team 10 of the probability of success for teams in the competition.

5.3.2 Results of competition

In Figure 20 (shown on the next page) see the results for the teams in the competition. The overall ranking, from first to last, was team 2, team 10, team 1, team 6, team 3, and finally team 5. The results indicate that teams 3, 5 and 6 performed poorly, that team 1 was an average performer, and that teams 2 and 10 were good-to-excellent performers.

The ranking (i.e., first to last) for bin number 10 was team 2, team 10, team 1, team 5, and teams 3 and 6. Note in Table 8 (shown on the next page) that the actual ranking by points scored in the Competition is shown as team 10, team 2, team 1, teams 5 and 6, and team 3 indicating some error in the weighting scheme (discussed in chapter 4) used to predict outcomes of events as well as display results of events. For bin number 5, the ranking (again first to last) was team 10, team 1, teams 3 and 6, team 2, and team 5. In bin number 1, the ranking (also first to last) was team 5, team 3, team 6, team 2, and teams 1 and 10.

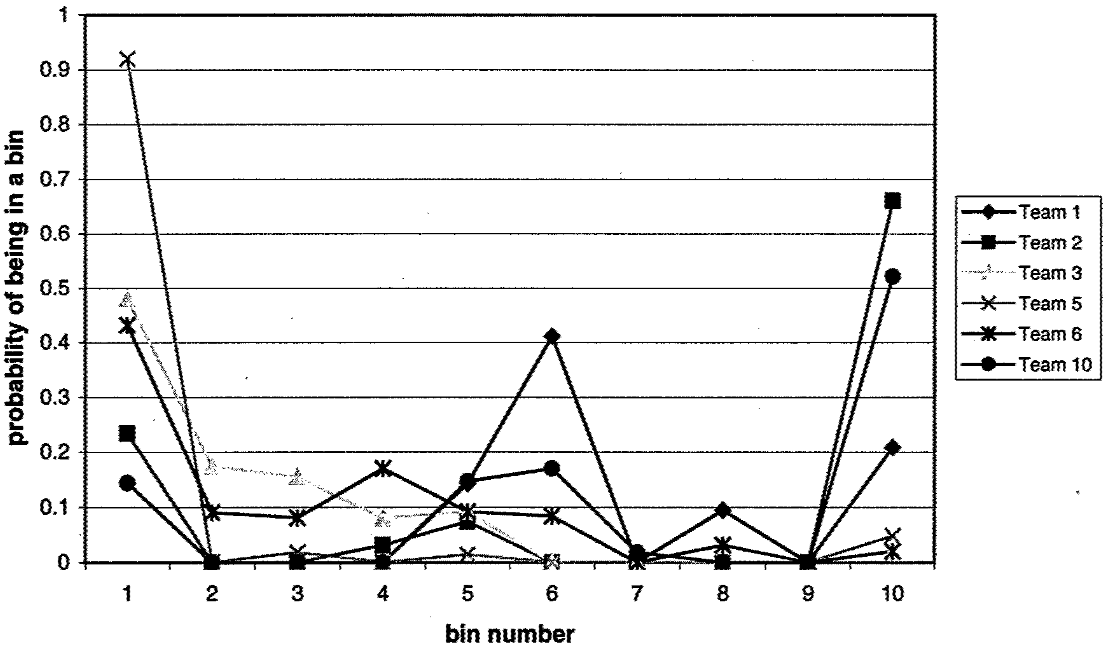

Figure 20: Competition results (participating teams).

Table 8: Competition results in terms of points awarded.

Team number	Points scored
10	12
2	11
1	4
5,6	1
3	0

5.3.3 A general description of faults of the robots in the Competition follows.

Some robots still exhibited difficulty in signaling the arrival at a target. Others, however, readily lifted a mechanical arm, played some music or made a whistle.

False signals occurred (i.e., signaling without being near a target).

Some robots stalled in a corner, at a wall, at scoring areas or at home. Some robots even stalled without having left home.

Numerous difficulties were exhibited in dealing with targets either in aspects of detection or of control. Some robots drove past or over targets or past score areas. Other robots pushed targets to score areas only to then push them out of those same score areas. Still other robots were able to collect numerous targets but were then unable to transport them to score areas.

Chapter 6

Bayesian approach to projecting success of robots in successive events

The value of Bayes' Theorem lies in the ability to predict the success of future events, in our case robot performance, in successive events. The requirements to use Bayes' Theorem are an initial data set predicting the success of the robots, in our case, and the results of that event (i.e., evidence). The participating teams provided the predictions of success for each of the participating teams by way of the questionnaires (i.e., evaluations projecting the success of the teams in Exhibition 1). The results of Exhibition 1 (and the subsequent events) were compiled by William Hardman and are described in detail in Section 3.2.

The two data sets just described (predictions of Exhibition 1 and the results of Exhibition 1) can then be combined, using Bayes' Theorem in producing the Bayesian-updated prediction for the next event (Exhibition 2). This Bayesian-updated prediction for Exhibition 2 can then be combined with the results, or evidence, from Exhibition 2, by using Bayes' Theorem for the second time to produce the Bayesian-updated prediction for the next event (i.e., the competition). The point is simple; the output of the first use of Bayes' Theorem provides input for the subsequent use of the theorem provided new evidence exists. If new evidence exists, then the second use of Bayes' Theorem can occur and provide a prediction of the next event. This process can be repeated as long as additional events occur and provide new evidence, each time.

6.1 Description of the Bayesian method

Bayes' Theorem (equation 2) is that which immediately follows with individual parts of the equations defined subsequently.

Equation 2: Bayes' Theorem [4]

$$P\left(\Theta = \theta_i, \varepsilon\right) = \frac{P\left(\varepsilon, \Theta = \theta_i\right) P\left(\Theta = \theta_i\right)}{\sum_{i=1}^{n} P\left(\varepsilon, \Theta = \theta_i\right) P\left(\Theta = \theta_i\right)}$$

$P\left(\varepsilon, \Theta = \theta_i\right)$ is the likelihood of experimental outcome ε if $\Theta = \theta_i$ (conditional probability)

$P\left(\Theta = \theta_i\right)$ the prior probability of $\Theta = \theta_i$ that is prior to the θ_i availability of the experimental information ε

$P\left(\Theta = \theta_i, \varepsilon\right)$ the posterior probability of $\Theta = \theta_i$ that is, the probability that has been revised in light of experimental outcome ε

Equation 2, or Bayes' Theorem, allows for consistent revision of performance expectations as new evidence becomes available that is relevant to the success of the robots in subsequent events. The prior probability described above is the evaluating (i.e., our participating) team's predictions of success (obtained from the questionnaires) for each of the participating teams for Exhibition 1. Given the prior probability and the results of the robot teams in Exhibition 1 (i.e., the actual evidence of the robots' performance), a Bayesian-updated prediction of success (i.e., the posterior probability) for Exhibition 2 is obtained using equation 1. Of note here is the fact that the average of the four non-participating teams results was used as the results, or evidence, for each of the participating teams as the participating teams were of similar capability though not

quite prepared to participate in Exhibition 1. Note also, that the results of the robots' performance, or evidence, correlates with the conditional probability in Bayes' Theorem. The above means that a posterior probability is obtained for each distribution obtained from the other teams. Of course, this means that the sum of the probabilities of all ten bins, for any one team, remains equal to unity.

Bayes' Theorem is used a second time to obtain a new posterior distribution based upon the team-specific results of Exhibition 2. The new posterior distribution is an indication of the projected success of the robots in the Competition (i.e., the next event). This requires that the posterior probability (the Bayesian-updated prediction of success for Exhibition 2) obtained above becomes the prior probability for this next use of Bayes' Theorem. That is, to obtain the current Bayesian-updated prediction of success in future events (i.e., the competition), the calculated probability (Bayesian-updated prediction of success for Exhibition 2) of success for the previous event must itself be updated, which is one way of looking at what Bayes' Theorem accomplishes.

Therefore, in using Bayes' Theorem a second time we take the new prior probability just described in conjunction with the results from Exhibition 2 (i.e., the evidence of the robots' performance) to obtain the posterior probability, applicable before the Competition. The posterior probability relevant to the Competition is the Bayesian-updated prediction of success for the Competition obtained by the same method as the posterior probability for Exhibition 2.

6.2 Error Calculation

Once Bayesian-updated predictions of success have been calculated for Exhibition 2 and the Competition, respectively, the next step in our analysis is to calculate root mean square errors (RMS) concerning these predictions of success and the actual results of the two events. The root mean squared error is obtained as shown in Equation 3. below.

Equation 3: Root Mean Squared Error (RMS) [5]

$$\mathrm{RMS} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} e_i^2}$$

In this analysis we take the sum of the difference squared between the actual results ($\hat{y}_i$ above) and the predicted results, y_i, then divide by the total number of terms, n (i.e., 10, corresponding to 10 bins), and finally take the square root in order to obtain the RMS.

Of value are two calculations of the RMS: that of the Bayesian predicted results and actual results and that of the team projections of success and actual results (both calculations done for Exhibition 2 and the Competition). These calculations can then be compared to those in the paragraph above to determine whether Bayesian updating more closely approximated the actual results or if the teams were able to predict more effectively the actual results. The method with the lower RMS indicates that that method has, in fact, predicted more effectively the actual results.

6.2.1 RMS error for Exhibition 2

The upper half of Table 9 and Figure 21 (on the next page) provide a tabulation of average RMS error and a graphical representation of RMS error between actual results and Bayesian projections for Exhibition 2, respectively. The tabulated data in Table 9 are a sum of the RMS errors as well as the average RMS error for all of the teams evaluated by teams 1, 3, 5, 6 and 10 (i.e., teams 1, 3, 5, 6 and 10 each evaluated teams 1, 2, 3, 5, 6 and 10). An RMS error was determined for each of the 6 teams evaluated; this was done for each of the 5 evaluating teams (i.e., a total of 30 calculations). An average RMS error, for each evaluating team, was obtained by taking a simple average of the RMS error values (6 total values) between a teams performance in the event (Exhibition 2) and the Bayesian-updated prediction of that team's success. Figure 21 offers the graphical representation of the data, which are reduced to average RMS error in the upper half of Table 9. Therefore, the evaluating team (for Exhibition 2) with the lowest average RMS error was the team whose Bayesian-updated predictions for Exhibition 2 most accurately predicted the results of exhibition 2. For Exhibition 2, a ranking of the teams was determined (lowest average RMS error to highest RMS average RMS error) as teams 1, 5, 6, 3 and 10 (upper half of Table 9).

Table 9: Average RMS error for Exhibition 2 (actual results vs. Bayesian and team projections). Each evaluating team (5 total) evaluated the 6 participating teams. From these evaluations, Bayesian-updated projections and team projections for Exhibition 2 were separately compared to the actual results of Exhibition 2. Six RMS errors resulted in each case. The 6 RMS errors for each case were summed and an average taken with the results displayed in this table.

Average RMS error between actual results and Bayesian projections for exhibition 2		
	error summed	average RMS error
Team 1	0.375	0.063
Team 3	1.142	0.190
Team 5	0.423	0.070
Team 6	0.671	0.112
Team 10	1.249	0.208
Average RMS error between actual results and team projections for exhibition 2		
	error summed	average RMS error
Team 1	1.545	0.257
Team 3	1.617	0.270
Team 5	1.523	0.254
Team 6	1.575	0.263
Team 10	0.737	0.123

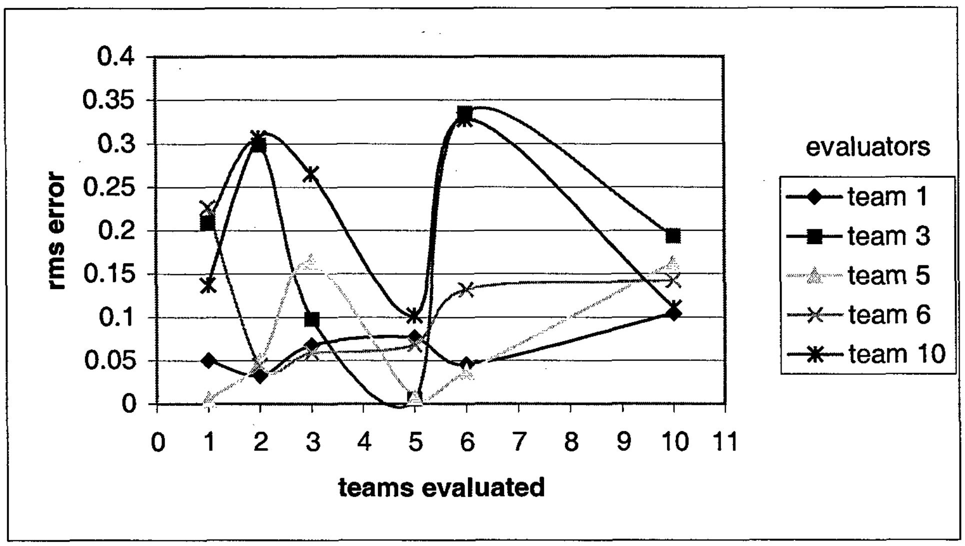

Figure 21: Each evaluating team (5 total) evaluated the 6 participating teams. From these evaluations, the Bayesian-updated projections for Exhibition 2 were compared to the actual results of Exhibition 2. Six RMS errors resulted. The 6 RMS errors associated with each of the evaluated teams is displayed in this figure for each of the 5 evaluating teams, called evaluators here. The data represented in this figure are reduced to average RMS errors for each evaluating team in the upper half of Table 9. rms error = $\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2}$, with actual results equal $\hat{y}_i$ and Bayesian projections equal y_i, and number of bins, n, equals 10.

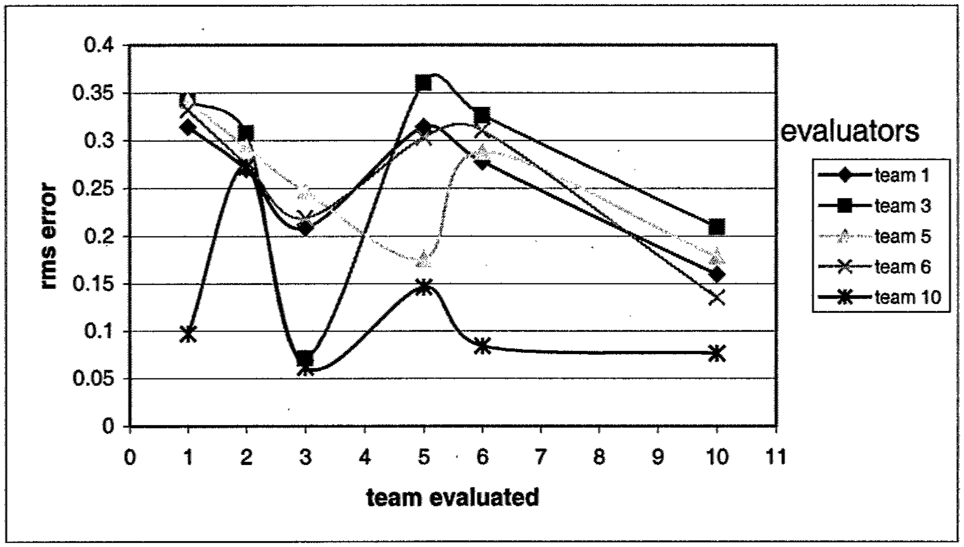

Figure 22: Each evaluating team (5 total) evaluated the 6 participating teams. From these evaluations, the team projections for Exhibition 2 were compared to the actual results of Exhibition 2. Six RMS errors resulted. The 6 RMS errors associated with each of the evaluated teams is displayed in this figure for each of the 5 evaluating teams, called evaluators here. The data represented in this figure are reduced to average RMS errors for each evaluating team in the lower half of Table 9. rms error = $\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2}$, with actual results equal $\hat{y}_i$ and team projections equal y_i, and number of bins, n, equals 10.

The lower half of Table 9 (previous page) and Figure 22 above provide a tabulation of average RMS error and a graphical representation of RMS error between actual results and team projections for Exhibition 2, respectively. The tabulated data in Table 9 represent a sum of the RMS errors as well as the average RMS error for all of the teams evaluated by teams 1, 3, 5, 6 and 10 (i.e., Teams 1, 3, 5, 6 and 10 each evaluated teams 1, 2, 3, 5, 6 and 10). An RMS error was determined for each of the 6 teams evaluated; this was done for each of the 5 evaluating teams (i.e., a total of 30 calculations). An average RMS error, for each evaluating team, was obtained by taking a simple average of the RMS error values (6 total values) between a team's performance

in the event (Exhibition 2) and the evaluating team's projections of that team's success. Figure 22 offers the graphical representation of the data, which are reduced to average RMS error in the lower half of Table 9. Therefore, the evaluating team (for Exhibition 2) with the lowest average RMS error was the team whose predictions for Exhibition 2 most accurately predicted the results of Exhibition 2. For Exhibition 2, a ranking of the teams was determined (lowest average RMS error to highest average RMS error) as teams10, 5, 1, 6, and 3 (lower half of Table 9). This ranking is similar to that described by Bayesian projection compared to actual results except Team 10 has moved from last above to first here and teams 1 and 5 above are in reversed order.

Of additional note in Table 9 is the fact that the average RMS error is lower on average by 0.152 for teams 1, 3, 5, and 6 (team 10 exceeds by 0.085) for the Bayesian projection versus actual results as compared to team projections versus actual results. This is an indication of the fact that team 10, to this point, is a fairly capable evaluator as it is able to perform evaluations with less RMS error than that obtained via Bayesian projections.

6.2.2 RMS error for Competition

The upper half of Table 10 and Figure 23 (both on page 70) provide a tabulation of average RMS errors and a graphical representation of RMS error between actual results and Bayesian projections for the Competition, respectively. The tabulated data in Table 10 represent a sum of the RMS errors as well as the average RMS error for all of the teams evaluated by teams 1, 3, 5, 6 and 10 (i.e., Teams 1, 3, 5, 6 and 10 each evaluated teams 1, 2, 3, 5, 6 and 10). An RMS error was determined for each of the 6

teams evaluated; this was done for each of the 5 evaluating teams (i.e., a total of 30 calculations). An average RMS error, for each evaluating team, was obtained by taking a simple average of the RMS error values (6 total values) between a teams performance in the event (the Competition) and the Bayesian-updated prediction of that team's success. Figure 23 offers the graphical representation of the data, which are reduced to average RMS errors in the upper half of Table 10. Therefore, the evaluating team (for the Competition) with the lowest average RMS error was the team whose Bayesian-updated predictions most accurately predicted the results of the competition. For the Competition, a ranking of the teams was determined (lowest to highest average RMS error) as 3, 5, 1 and 6 (equal average RMS error), and 10. This ranking shows consistency only for teams 5, 6 and 10. That is, the data for Exhibition 2 (upper half of Table 9) put these teams in the same positions as seen here.

Table 10: Average RMS error for Competition (actual results vs. Bayesian and team projections). Each evaluating team (5 total) evaluated the 6 participating teams. From these evaluations, Bayesian-updated projections and team projections for the Competition were separately compared to the actual results of the Competition. Six RMS errors resulted in each case. The 6 RMS errors for each case were summed and an average taken with the results displayed in this table.

Average RMS error between actual results and Bayesian projections for competition		
	error summed	average RMS error
Team 1	1.328	0.221
Team 3	1.069	0.178
Team 5	1.319	0.220
Team 6	1.326	0.221
Team 10	1.355	0.226
Average RMS error between actual results and team projections for competition		
	error summed	average RMS error
Team 1	1.173	0.196
Team 3	1.294	0.216
Team 5	1.153	0.192
Team 6	1.064	0.177
Team 10	1.101	0.183

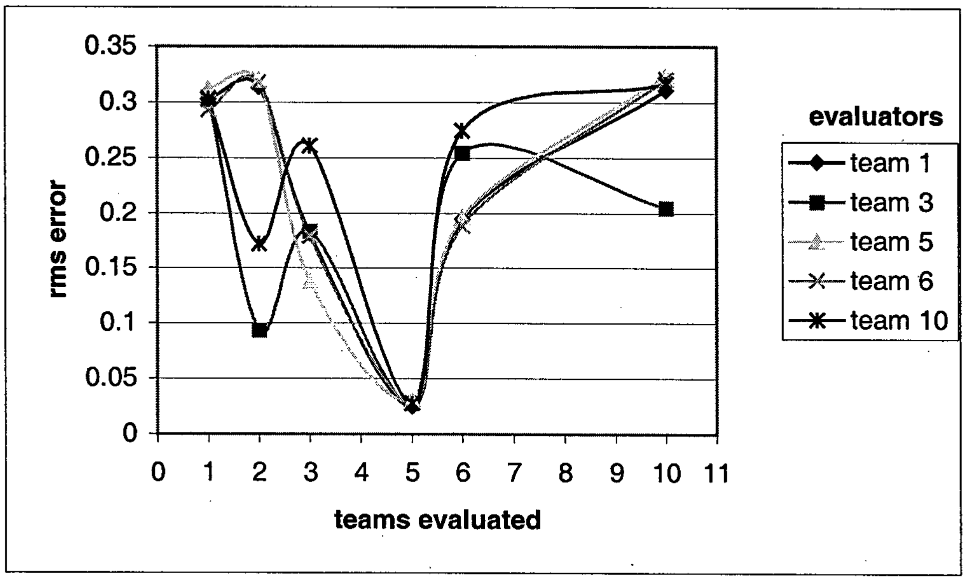

Figure 23: Each evaluating team (5 total) evaluated the 6 participating teams. From these evaluations, the Bayesian-updated projections for the Competition were compared to the actual results of the Competition. Six RMS errors resulted. The 6 RMS errors associated with each of the evaluated teams is displayed in this figure for each of the 5 evaluating teams, called evaluators here. The data represented in this figure are reduced to average RMS errors for each evaluating team in the upper half of Table 10. rms error = $\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}$ **, with actual results equal** $\hat{y}_i$ **and Bayesian projections equal** y_i**, and number of bins, n, equals 10.**

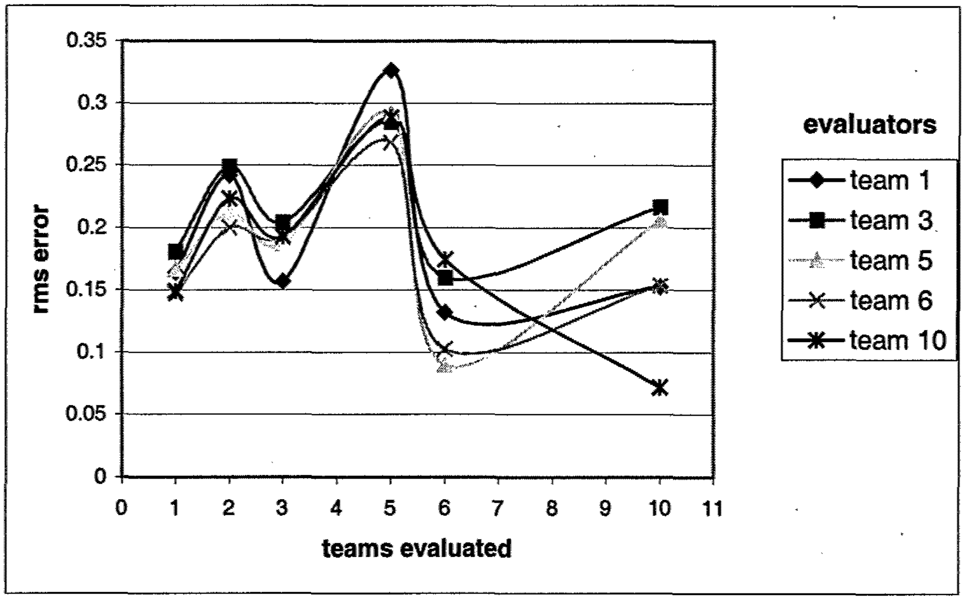

Figure 24: Each evaluating team (5 total) evaluated the 6 participating teams. From these evaluations, the team projections for the Competition were compared to the actual results of the Competition. Six RMS errors resulted. The 6 RMS errors associated with each of the evaluated teams is displayed in this figure for each of the 5 evaluating teams, called evaluators here. The data represented in this figure are reduced to average RMS errors for each evaluating team in the lower half of Table 10. rms error = $\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2}$, with actual results equal $\hat{y}_i$ and team projections equal y_i, and number of bins, n, equals 10.

The lower half of Table 10 and Figure 24 above provide a tabulation of average RMS error and a graphical representation of RMS error between actual results and team projections for the Competition, respectively. The tabulated data in Table 10 represent a sum of the RMS errors as well as the average RMS error for all of the teams evaluated by teams 1, 3, 5, 6 and 10 (i.e., teams 1, 3, 5, 6 and 10 each evaluated Teams 1, 2, 3, 5, 6 and 10). An RMS error was determined for each of the 6 teams evaluated; this was done for each of the 5 evaluating teams (i.e., a total of 30 calculations). An average RMS error, for each evaluating team, was obtained by taking a simple average of the RMS error values (6 total values) between a team's performance in the event (the Competition) and

the evaluating team's prediction of that team's success. Figure 24 offers the graphical representation of the data, which are reduced to average RMS errors in the lower half of Table 10. Therefore, the evaluating team (for the Competition) with the lowest average RMS error was the team whose predictions for the Competition most accurately predicted the results of the Competition. For the Competition, a ranking of the teams was determined (lowest to highest average RMS error) as teams 6, 10, 5, 1 and 3 (lower half of Table 10). This ranking is consistent for all teams with the exception of team 6, which moves from second to last position (lower half of Table 9) to the first position here. Of additional note in Table 10 is the fact that the average RMS error is lower on average by 0.035 for teams 1, 5, 6 and 10 (team 3 exceeds by 0.038) for the actual results versus team projections as compared to Bayesian projection versus actual results. This means that the team projections (compared to actual results) for all but team 3 were marginally better than the Bayesian projections.

Chapter 7

Overall observations regarding the contests

In Appendix F are found the team rankings as evaluated by each of the five evaluating teams (i.e., this is an expected bin approach to ranking the teams as per Equation of Chapter 5) as well as the results of each of the three events. In addition, the Bayesian-updated projections of success for Exhibition 2 and the Competition (i.e., Posterior for Exhibition 2 and Posterior for the Competition, respectively) are, likewise, shown in Appendix F. This Appendix, like all others, is for reference and is simply Table 7 with two additional columns (i.e., the posteriors for Exhibition 2 and the Competition). The tables and figures in Chapter 6 are those, which are used in drawing some conclusions here.

7.1 Exhibition 1 evaluation problems

Because teams 1, 2, 3, 5, 6 and 10 were unprepared for Exhibition 1, this created the necessity of using data, which approximated the capability of these teams. These data were taken from the non-participating teams (4, 8, 9, and 11) as shown in Figure 7. These data were fairly representative of the actual capability of the participating teams as evidenced by the RMS error between Bayesian predictions of success in events as compared to the results of those events. Though the data were representative, it would have been preferable to have data from the participating teams as the starting point for these updating studies. If a fourth event had occurred (i.e., an event following the Competition), then the Exhibition 1 data could have either been kept as is and updated

three times using Bayes' Theorem or thrown out. The Exhibition 1 data would only have been thrown out provided Exhibition 2 data updated twice to predict the success of the teams in the hypothetical fourth event gave better results (i.e., lower RMS errors).

Team 6 provided no evaluations for Exhibition 1. This was problematic since no Bayesian-updated predictions of Exhibition 2 or the Competition could be calculated. Instead of disregarding team 6 evaluations for Exhibition 2 and the Competition, the team 6 evaluations for Exhibition 2 were substituted for the, non-existent, evaluations for Exhibition 1. We believed, at first, that this fact could have invalidated the Bayesian-updated team 6 evaluations, which predicted the success or failure of the teams in Exhibition 2. However, upon reviewing the average RMS errors for Exhibition 2 as are shown in Table 9, we note that the average RMS error for team 6 (for Exhibition 2) in the case of the Bayesian projections is third lowest (of five). In effect, using the team 6 evaluations done prior to Exhibition 2 as input (i.e., in place of the non-existent team 6 Exhibition 1 evaluations) for the first use of Bayes' Theorem (in Bayesian updating) did not adversely affect the team 6 Bayesian projections for Exhibition 2.

If the robots had been more capable of performing the required tasks in Exhibition 1, then a more rapid determination of the most useful questions or questions that should have been included but were not would have occurred. As it was, questions that should have been tossed out like "Does the robot damage the playing surface?" remained through the termination of the Exhibition 2. Earlier identification of useless questions and search for questions of greater value in identifying distinguishing characteristics between robots would provide for a more effective questionnaire which would allow for greater success in predicting success of the teams in future events.

7.2 Accuracy and consistency changes (from Exhibition 2 to the Competition) in evaluations performed by each of the five evaluating teams

The team rankings, derived from the Bayesian-updated evaluation average RMS errors, changed from the Exhibition 2 to the Competition, as is shown in Table 11 and is discussed below. These rankings are, in effect, an indication of how well the teams performed as evaluators (i.e., the evaluating team's accuracy in evaluation); the higher the rank the more accurate the team was as an evaluator. If no new information was available between the end of Exhibition 2 and the Competition, the rankings should have remained consistent. However, the rankings, in fact, take into account the effect of new data between the two events.

Table 11: Shows the ranking of the evaluating teams based on average RMS error in evaluating all 6 of the evaluated teams for both Exhibition 2 and the Competition. For Exhibition 2 and the Competition, team rankings based on average RMS errors were derived for the Bayesian-updated team evaluations (or posteriors) and the team evaluations (i.e., not Bayesian-updated). The ranking of the teams is from the lowest average RMS error (best) shown at the top of a column to the highest average RMS error (worst) shown at the bottom of the column.

	Exhibition 2		**Competition**	
rank	evaluating teams listed in order of lowest RMS error in Bayesian-updated team evaluations	evaluating teams listed in order of lowest RMS error in team evaluations	evaluating teams listed in order of lowest RMS error in Bayesian-updated team evaluations	evaluating teams listed in order of lowest RMS error in team evaluations
1	1	10	3	6
2	5	5	5	10
3	6	1	1,6	5
4	3	6	10	1
5	10	3		3

Note also that it is preferable to be the team ranked low in Exhibition 2 and ranked significantly higher in the Competition, as it is indicative of improvement in evaluation ability, unless of course a team was ranked high throughout the competition. In which case, that team would have evaluated exceptionally well from the beginning.

The team rankings from the Bayesian-updated evaluation average RMS errors are as follows. Team 1 went from the best evaluator, or most accurate, to third of five. Team 3 improved from fourth to first of five. Teams 5, 6 and 10 remained consistent at second of five, third of five, and fifth of five (or least accurate), respectively.

The team rankings, derived from the team evaluation (i.e., not Bayesian-updated) average RMS errors, changed as shown in Table 11 and as is discussed below. These rankings, just as those discussed in the previous paragraph, are an indication of the evaluating team's accuracy in evaluation. The difference here is that no Bayesian updating has been performed on this data set. Team 1 went from the third best evaluator to fourth of five. Team 3 remained consistent at fifth of five. Team 5 slipped from second to third of five. Team 6 improved from fourth to first of five. Team 10 was downgraded from first to second of five.

7.3 Accuracy and consistency changes (from Exhibition 2 to the Competition) in evaluations of the six evaluated teams

7.3.1 Teams evaluated well as determined via the Bayesian-updated team evaluation average RMS errors

Here an arbitrary assumption was made that an RMS error of 0.1 (or less) was considered a good evaluation. For Exhibition 2, the Bayesian-updated team evaluation

RMS errors (for Exhibition 2) of Figure 21 show that the following teams were evaluated well and by which teams those evaluations occurred. Teams 1 and 5 were successfully able to evaluate team 1. Teams 1, 5 and 6 successfully evaluated team 2. Teams 1, 3 and 6 evaluated well team 3. All teams (teams 1, 3, 5, 6 and 10) evaluated well team 5. Teams 1 and 5 successfully evaluated team 6. Teams 1 and 10 evaluated successfully team 10.

If instead we now use an RMS error of 0.2 in Figure 21 to define a good evaluation for the Bayesian-updated team evaluations of Exhibition 2, additional teams are added to the list of good evaluators. Teams 3 and 10 successfully evaluated team 1. Team 5 performed well in evaluating team 3. Team 6 evaluated itself well. Teams 3, 5, and 6 successfully evaluated team 10.

Here the assumption was made that an RMS error of 0.1 (or less) was considered a good evaluation. For the Competition, the Bayesian-updated team evaluation RMS errors of Figure 23 show that the following teams were evaluated well and by which teams those evaluations occurred. Team 3 successfully evaluated team 2. All teams (teams 1, 3, 5, 6 and 10) evaluated team 5 well. The only consistency between Exhibition 2 and the Competition was the fact that all teams were able to evaluate team 5 well. This means that each of the evaluating teams was consistently able to identify a relatively mediocre (in performance) team (see Table 8).

If instead we now use an RMS error of 0.2 in Figure 23 to define a good evaluation for the Bayesian-updated team evaluations of the Competition, again additional teams are added to the list of good evaluators. Team 10 performed well in its

evaluation of team 2. Teams 1,3, 5, and 6 successfully evaluated team 3. Teams 1, 5, and 6 evaluated team 6 well. Team 3 successfully evaluated team 10.

In the case of 0.2 RMS errors for the Bayesian-updated team evaluations, consistency between Exhibition 2 and the Competition was as follows below. Teams 1, 3, 5, and 6 successfully evaluated team 3 for both events. All teams evaluated well team 5 in both events. Teams 1, 5, and 6 successfully evaluated team 6 for the two events. Team 3 successfully evaluated team 10 in both events. Of note is the fact that teams 1, 5, and 6 evaluated effectively and consistently teams 3, 5, and 6 (all mediocre teams in performance). Team 3, on the other hand, was able to evaluate effectively and consistently teams 3 and 5 (both mediocre teams in performance) and team 10 (a high performer).

In addition we should consider more closely, the Bayesian-updated team 3 evaluation of team 2 (an excellent performer in the competition). If we discount the fact that team 3 had only the fourth lowest RMS error in its Bayesian-updated evaluation of team 2 prior to Exhibition 2 (see Figure 21), and instead look at the team 3 Bayesian-updated evaluation of team 2 prior to the Competition (see Figure 23), we see that team 3 had the best showing with an RMS error of 0.09 and team 10 was second with an RMS error of 0.17. It is easier to believe that a team is a good evaluator if it is capable of evaluating well both good and poor performers. In this case the lack of consistency between Exhibition 2 and the Competition should be ignored, as the Bayesian-updated team 3 evaluation of team 2 was extremely accurate in the Competition, the event that counted. This makes team 3 our choice for the best evaluator in this instance.

7.3.2 Teams evaluated well as determined via the team evaluation average RMS errors

Again, the assumption was made that an RMS error of 0.1 (or less) was considered a good evaluation. For Exhibition 2, the team evaluation average RMS errors of Figure 22 show that the following teams were evaluated well and by which teams those evaluations occurred. Team 10 was successfully able to evaluate teams 1, 3, 6 and 10. Team 3 successfully evaluated team 3.

If instead we now use an RMS error of 0.2 in Figure 22 to define a good evaluation for the team evaluations of Exhibition 2, additional teams are added to the list of good evaluators. Team 1 successfully evaluated team 3. Teams 5 and 10 performed well in evaluating team 5. Teams 1, 3, 5, and 6 successfully evaluated team 10.

Here the assumption was made that an RMS error of 0.1 (or less) was considered a good evaluation. For the Competition, the team evaluation RMS errors of Figure 24 show that the following teams were evaluated well and by which teams those evaluations occurred. Teams 5 and 6 successfully evaluated team 6. Team 10 evaluated itself well. The only consistency between Exhibition 2 and the Competition was the fact that Team 10 was able to successfully evaluate itself over the two events.

If instead we now use an RMS error of 0.2 in Figure 24 to define a good evaluation for the team evaluations of the Competition, additional teams are added to the list of good evaluators. All teams successfully evaluated team 1. Team 6 performed well in evaluating team 2. All teams successfully evaluated team 3. Teams 1, 3, and 10 successfully evaluated team 6. Teams 1, 5, and 6 evaluated team 10 well.

In the case of 0.2 RMS errors for the team evaluations, consistency between Exhibition 2 and the Competition was as follows below. Team 10 successfully evaluated team 1 in both events. Teams 3 and 10 successfully evaluated team 3 for both events. Teams 10 successfully evaluated team 6 for the two events. Teams 1, 5, 6 and 10 successfully evaluated team 10 in both events. This says that team 10 was consistently adept at evaluating four teams (1, 3, 6 and 10), which were three poor performers and itself (a high performer). Of note, however, is the fact that team 10, a good evaluator as well as a high performer in the events, did not evaluate its stiffest competition, team 2, effectively (i.e., RMS errors slightly greater than 0.2, specifically 0.28 for Exhibition 2 and 0.23 for the Competition). However, in reviewing Figures 22 and 24, we can see that team 10 was one of the best (if not the best) evaluators of team 2, which may imply that the (arbitrary) 0.2 RMS error cutoff in determining if an evaluator was proficient may in fact be too low. Of additional note, is the fact that three mediocre performers (teams 1, 5 and 6) consistently evaluated team 10 (a high performer) effectively. Again, it is easier to believe that a team is a good evaluator if it is capable of evaluating well both good and poor performers. In this instance we choose team 10 as the best evaluator.

Chapter 8

The link between the nuclear research and development example and the robot design contest

8.1 The nuclear research and development example compared to the robot design contest

Twenty-seven criteria were used in determining which nuclear power plant concepts were well suited to a given mission. The 27 criteria were used in creating three major goals (sustainability, safety & reliability, and economics). That is, the score assigned in each of the 27 areas was rolled into 3 major scores called goals in this study. In the case of the robot design course, 17 criteria were used in the creation, likewise, of three major goals (control, control/score, and score).

As is noted previously, problems existed in the case of the nuclear example, which do not exist in the robot design course. First, the work conducted was purely hypothetical (i.e., no power plants have yet been built) and therefore no operational comparisons between plants could be made, let alone multiple operational comparisons to provide a fairly exact assessment of plant capability. Second, inherent bias existed because the concept designers evaluated their own designs.

The robot design course quite readily handled the above two problems from the nuclear example. First, the work was carried to practical realization; the robots were constructed and multiple (three) operational comparisons occurred allowing for a

rigorous assessment of each robot's capabilities. Second, the concept designers evaluated their robot as well as those of the competition. Finally, these evaluations were then compared to the operational results of the events (exhibitions and competition). The fact that the evaluations must stand next to the results of the events lessens the bias a robot designer might apply to his own design especially in subsequent evaluations.

8.2 Improving the nuclear research and development process example based on work in robot design course

The fact remains, however, that a more robust evaluation method for the nuclear example (i.e., the impetus for this study) is desired. The methodology used in the robot design course can be incorporated into the nuclear example. Surveys (i.e., questionnaires) can be conducted by independent evaluators, from which, probability mass functions concerning the performance variables can be derived. The probability mass functions can be used to eliminate those designs not scoring above some threshold level. Further, more in-depth evaluations can follow to eliminate other concepts until a select group of a few concepts remains. Obviously, preliminary designs will become more concrete as designs are selected for further evaluation.

How could the surveys be conducted? An independent evaluation team could evaluate all of the prospective designs in order to remove the biases of individual designers toward their own designs. The evaluators' independence from any of the competing designs would help to ensure that the evaluation is fair. The evaluation team could conduct a survey (i.e., a series of questions) designed to provide a probability mass function predicting the success of a given design. The concept designers will have to

justify (to the satisfaction of the evaluators) answers to any question and the concept designers will not be privy to the scores assigned. This evaluation process will be done in stages with the threshold being raised for each subsequent evaluation. This process will occur until a few surviving concepts remain.

Following the down-selection to the few surviving concepts, construction of small-scale prototypes of the remaining concepts could occur. Operational tests could then be conducted. For each operational test, after the first test, Bayesian updating could be used to predict the success of each of the concepts as done with the robot designs. Bayesian updating could reduce the number (by one) of operational tests required as it can permit prediction of the success of a design in the next event. Doing this could reduce the number of operational tests required for the remaining two or three concepts, and thus, the costs involved before final selection of the winning concept.

Chapter 9

Conclusions

9.1 Recommendations for future work in improving evaluation methods

9.1.1 Methodology used by teams in assigning a score to each of the teams

The methodology used by each of the teams in assigning a score to the teams that it evaluated varied from team to team. The methodology specific to each team is shown in Appendix G. The methods, in general are described below.

1. The first step is to read the Wiki journal [3] daily to learn the best practices of other teams and avoid problems encountered by other teams. The Wiki Journal, found on the MASLAB website [3], is a journal in which the design teams update their daily progress. In addition, speak directly with the teams to identify the best practices and the problems encountered that may have been left out of the journal entries.

2. The second step is to observe other robot design teams in the lab environment and make comparisons to own design and progress. Keep track of how much effort is put into the robot as an indication of expected success in the upcoming contests. Use performance in previous events as evidence in evaluations of projected success in future events.

3. The third step is to question whether other teams often lack direction or consistently know what they are doing? Were other teams able to build sensors on their

own? The sensors are fairly sophisticated and the ability to readily employ them in the robot design was an indicator, potentially, of overall design sophistication and an indicator of success in the competitive events. Were teams using water jetting (i.e., a sophisticated manufacturing method) for the robot chassis? If a team was using such a method in the manufacture of the robot chassis, their robot was likely to be fairly sophisticated overall and excellent performance, thus, was expected. Were teams consistently making last minute preparations prior to the contests? If so, poor performance would likely be the result.

9.1.2 What would have made the evaluation process easier for the teams?

Each of the teams provided information regarding what would have made the evaluation process easier for them. The team specific information can be found in Appendix H. A general description of what would have simplified the process follows below.

Provide (to the evaluators) questionnaires better suited to each event such that, not only a current, but also a fairly exact assessment of each robot capability could be obtained by filling out such surveys. The capabilities of the robots were not equal in each of the contests. Additionally full capability was not achieved until the final contest, the Competition. These facts meant that some of the questions asked in the questionnaire did not apply until Exhibition 2 and possibly until the Competition.

Spend more time in filling out the questionnaire. The contests themselves offered the best evidence of a robot's capability but the first evaluation was based largely on observations made in the lab prior to Exhibition 1, making for a difficult first evaluation. This meant that more time must be spent in observing in the lab in order to

make the first evaluation. Follow-on evaluations must also be based more on the observations in the lab since the lab is the second most important opportunity (next to the competitive events themselves) to obtain the data used in evaluating each of the teams. Additional time could have been spent in discussing problems as well as best practices and plans of attack with each of the competing teams and in reading the Wiki journal [3].

Be consistent in filling out the questionnaires. That is, having previous questionnaire(s) available could assist in maintaining greater consistency in answering the questionnaires.

9.2 What we learned from the robot contests

9.2.1 Discrepancy occurred in determination of which team was the best evaluator.

Combining the evidence and conclusions in Chapter 7, we can now say that two teams were declared the best evaluators, but in two different circumstances (Bayesian-updated versus non-Bayesian updated team evaluations). If we adhere to the results obtained from Bayesian-updated team evaluations, then team 3 is the clear winner. Team 3 consistently and effectively evaluated teams 3, 5 and 10 where teams 3 and 5 were poor performers while team 10 was a high performer. However, if we adhere to the results obtained from team evaluations (i.e., non-Bayesian-updated), then team 10 is easily the best evaluator. Team 10 consistently and effectively evaluated teams 1, 2, 3, 6 and 10 where teams 1, 3, and 6 were poor performers and teams 2 and 10 high performers. Considering the above facts, then, the best evaluator overall is team 10 for the case of the non-Bayesian updated team evaluations.

Section 9.2.1 details the methods used by team 10, which gives some insight into why team 10 was able to perform so well as evaluators. In Section 9.2.1 we note that team 10 had the most complete and sophisticated explanation of what should be done to effectively evaluate the robot teams.

Why do we see a discrepancy between the Bayesian-updated choice of best evaluator and non-Bayesian-updated choice? Potentially many reasons exist. The evaluations completed prior to Exhibition 1 were completed with minimal information (i.e., no previous events had occurred by which to judge the robots). Therefore, the teams were only able to evaluate the other teams by means of observations made in the lab and via discussions with the other teams. Combine this fact with the fact that the results of Exhibition 1 were estimated because the participating teams were not prepared for Exhibition 1 and the Bayesian-updated team evaluations projecting the success of the teams in Exhibition 2 may not have made these data the most desirable. In other words, the evaluators went through two steep learning curves in attempting to meet the requirements of Exhibition 1(i.e., becoming technically competent) as well as attempting to become proficient evaluator of their peers.

Additionally, the teams were only minimally capable to compete by the time of Exhibition 1 and for that matter, that of Exhibition 2 as well. It would have been preferable to have had a fourth event to allow slightly less advanced teams to make the changes that would, perhaps, have changed their robot from a poor performer to a good or even high performer.

Examples of the previous statement include the non-official results of robots competing immediately following the Competition. The robot teams (not all teams

participated) made minor improvements after that team's performance in the Competition and were allowed at the conclusion of the competition to compete, unofficially. Team 3 scored 4 points (They had scored 0 points in the Competition.). Team 4 (one of the teams not participating in our project) scored 8 points (They had scored 0 points in the Competition.). Team 9 (another team not participating in our project) scored 13 points, which is one more point than the winning team (team 10) scored in the Competition.

In other words, robot progression from one event to another was more of an exponential change rather than a linear change. This fact makes evaluation harder for the novice evaluator, although team 10 was able to perform extremely well.

9.2.2 Methodology of the best evaluator (team 10)

As part of this work it was important to understand better how qualified experts evaluate technological concepts. The methodology used by team 10 in evaluating the robot teams prior to the three events follows below. Note that team 10 had the most complete and sophisticated (among the participating teams) idea or explanation of what should be done to effectively evaluate the robot teams.

The method that team 10 used included the following steps.

1. Observe the robot teams in the lab environment. Observe whether other teams often lacked direction or consistently know what they are doing? Were other teams able to build sensors on their own? The sensors are fairly sophisticated and the ability to readily employ them in the robot design was an indicator, potentially, of overall design sophistication and an indicator of success in the competitive events. Were teams using water jetting (i.e., a sophisticated manufacturing method) for the robot chassis? If a team

was using such a method in the manufacture of the robot chassis, their robot was likely to be fairly sophisticated overall and excellent performance, thus, was expected. Were teams consistently making last minute preparations prior to the contests? If so, poor performance would likely be the result.

2. Observe of team performance in the competitive events. This is the most obvious and, likely, the most useful means by which the teams should be evaluated. Performance in previous events is good evidence of the projected success of the teams in future events.

3. Identify problems and best practices of the other teams. This can be accomplished by reading the Wiki Journal [3], an online source of information, and discussing these concerns with the other teams. When team 10 was interviewed for their response as to what their method of evaluation consisted of, they did not state that they had used the Wiki Journal. Although team 10 did not, specifically, indicate that they had referred to the Wiki Journal, we assume that they, in fact, did so as did all the other teams. In addition to using the Wiki Journal, we likewise assume that team 10 also spoke with the other teams to identify problems and best practices that may not have been included in the Wiki Journal.

9.2.3 Success of the probabilistic approach to predicting the performance of technological concepts (i.e., the robots)

In addition to understanding better how qualified experts evaluated technological concepts, we wanted to determine if a probabilistically formulated method of integrating knowledge of various performance attributes provides better understanding of the likely performance of a technological concept. This was, in fact accomplished. Chapter 3

describes how the data (projections of the team's success) were obtained via questionnaires. Chapter 4 details how weightings were assigned to the data (i.e., how probabilities were assigned identifying the likelihood of a team's success in specific performance areas). Chapter 5 compares the data to the results of the three events. Chapter 6 shows the errors between the team projections of success and the actual results. Chapter 7 shows that, although some teams were better evaluators than others, successful evaluations of the teams occurred. Therefore, a probabilistically formulated method of integrating knowledge of various performance attributes can provide a better understanding of the likely performance of a technological concept, in our case robot performance.

9.3 Lessons which can be extended to the broader concern of how companies allocate R & D funding, to include the acquisition of a new Naval vessel or targeting technology investments in the drug delivery process.

Whatever the various options are for which R & D funding can be allocated (e.g., the newest Naval vessel or technology investment in the drug delivery process), a thorough and sophisticated evaluation of those options should occur. The overwhelming success of team 10 in evaluating effectively five of six participating robot teams was largely due to the thorough and sophisticated process used in evaluating the robot teams. Yes, this could go without saying and, yet, the other four evaluators (teams 1, 3, 5 and 6) had lesser degrees (some significantly) of success in their evaluations.

Both the Naval ship acquisition process and the technology investment in the drug delivery process specified phases that must be passed through to effectively obtain the

best ship and the technology which will best be used to increase research productivity, respectively. Within these phases used to ultimately select and produce the desired outcomes, lies the opportunity to employ a probabilistically-formulated method of integrating knowledge of various performance attributes that can provide a better understanding of the likely performance of the technological concept, in our two examples the ability of the ship to meet the mission need requirements and the ability to increase research productivity (ship acquisition and targeting technology investment in the drug delivery process, respectively).

First, surveys should be created by the individuals providing the requirements that must be met by the new ship design or new technology used to improve research productivity. Surveys can be completed at significant milestones or phases in the process. The surveys at the start of a specific milestone can be compared to the results obtained at the end of that phase. Bayesian updating (or the use of Bayes theorem) can be used to predict the future success of the ship design or the new technology at the end of the next phase. If only one design or new technology is under consideration, the Bayesian-updated prediction may very well produce the data used to determine if the process will continue or if the results do not justify continued support. If multiple designs or technologies are considered, the Bayesian-updated prediction may well be used to down-select one, or more, of the competing designs or technologies at any phase in the selection process.

Bayesian updating, in effect, can be used to limit the resources allocated to specific designs or new technologies by shortening the selection process. The Bayesian updating forecasts the success of the design or new technology in the next phase. If that

forecast does not meet the threshold requirement for that phase, then the funding for that design or new technology, not meeting the requirements, would be terminated. Only those designs or new technologies meeting the threshold requirement for that stage of the process would be allowed additional funding.

List of References

[1] Federal News Service, April 29, 2002 Monday, Department Defense Briefing, Special Defense Department Briefing, Subject: The Navy's Downselect Decision for the Lead Design Agent for the DD (X) Ship Program, Location: Defense Department Briefing Room, Arlington, Virginia.

[2] Research Technology Management, March 2003, Vol. 46, No. 2; Pg. 47; ISSN: 0895-6308, 03746001, Linking technological change to business needs: roadmapping offers a systematic approach for the pharmaceutical-biotechnology industry to target technology investments in the drug delivery process.

[3] Mobile Autonomous Systems Laboratory (Maslab/6.186), January 2003 <http://maslab.lcs.mit.edu>.

[4] Alfredo H-S. Ang and Wilson H. Tang, *Probability Concepts in Engineering Planning and Design, Volume I – Basic Principles*. New York, Chichester, Brisbane, Toronto, Singapore: John Wiley & Sons, Inc., pp. 329-336, 1975.

[5] Resa Corporation and Licensors, Descriptive Statistics, Root Mean Squared Error, 2002-2003 <http://www.xycoon.com/lsRoot%20Mean%20Sq.%20Error.htm>.

Appendices

Intentionally left blank

Appendix A: Exhibitions 1 & 2 questionnaire

The questionnaire used in Exhibitions 1 and 2 is presented below. The questionnaire shows the questions used to project a robot's potential prior to an event and to identify a robot's performance in the events. Minor changes were made between the questionnaire used for the Exhibitions and that used for the Competition (Appendix D). These changes are discussed in the text.

Questionnaire to be filled out by students of 6.186 who are doing the special project also.

Fill out one for your team and then one for each of the participating teams. The participating teams are 1, 2, 3, 5, 6, 10
Team members should split the work so only 2 or 3 questionnaires maximum will be filled by any one person.

Return to a staff member or Bill Hardman prior to the exhibition.

Team filling out the questionnaire ______ 1 2 3 5 6 10
(circle team number for which questionnaire filled out)

1 robot requires more than 60 seconds calibration — average performance (requires between 1 sec< time< 60 sec) — requires no calibration

1 2 3 4 5 6 7 8 9 10

select range

select peak

2 robot begins computing or starts prior to 10 second required delay — average performance (moves starting between 30 and 10 sec) — begins moving at exactly 10 seconds

1 2 3 4 5 6 7 8 9 10

select range

select peak

3	robot damages playing surface (includes leaving marks)				average performance				robot does not damage playing surface (leaves no marks)		
	1	2	3	4	5	6	7	8	9	10	

select range
select peak

4	does not arrive (i.e, within 4") at any specified waypoints				average performance				arrives (i.e., within 4") at all specified waypoints		
	1	2	3	4	5	6	7	8	9	10	

select range
select peak

5	does not perform waypoint signal after arriving (i.e, within 4") at any specified waypoints				average performance				performs waypoint signal every time within 4" of a waypoint		
	1	2	3	4	5	6	7	8	9	10	

select range
select peak

6	retrieved and placed in yellow scoring locations far less than average # of targets				average performance				retrieved and placed in yellow areas greater than average # of targets		
	1	2	3	4	5	6	7	8	9	10	

select range
select peak

7	robot unable to get to yellow scoring area				average performance				robot always arrives at yellow scoring area	
	1	2	3	4	5	6	7	8	9	10

select range
select peak

8	robot unable to detach target to deposit in yellow scoring area				average performance				robot always detaches target for deposit in scoring area	
	1	2	3	4	5	6	7	8	9	10

select range
select peak

9	retrieved and placed in home scoring location far less than average # of targets				average performance				retrieved and placed in home scoring area greater than average # of targets	
	1	2	3	4	5	6	7	8	9	10

select range
select peak

10	when robot fails to grasp target not smart enough to try again				average performance				when robot fails to grasp target is smart enough to try again	
	1	2	3	4	5	6	7	8	9	10

select range
select peak

11	*collides with all large* objects placed in its path				average performance				collides with no objects placed in its path	
	1	2	3	4	5	6	7	8	9	10

select range
select peak

12	robot unreliable in detecting scoring areas				average performance				robot 100% reliable in detecting scoring areas	
	1	2	3	4	5	6	7	8	9	10

select range
select peak

13	robot unreliable in detecting the targets				average performance				robot 100% reliable in detecting the targets	
	1	2	3	4	5	6	7	8	9	10

select range
select peak

14	robot has long processing time in detecting scoring areas and/or targets				average performance				robot has short processing time in detecting scoring areas and/or targets	
	1	2	3	4	5	6	7	8	9	10

select range
select peak

15 small percentage of playing field explored				average performance				large percentage of playing field explored		
	1	2	3	4	5	6	7	8	9	10

select range
select peak

16 robot does not arrive home in 4 minutes allowed				average performance				robot arrives home with time to spare		
	1	2	3	4	5	6	7	8	9	10

select range
select peak

17 robot cannot operate reliably on battery power				average performance				robot operates reliably on battery power		
	1	2	3	4	5	6	7	8	9	10

select range
select peak

Appendix B: Data collected from Exhibition 1

Evaluations were performed by teams 1, 3, 5, and 10 for Exhibition 1. All of the following applies to these teams but not to team 6, which did not provide evaluations for Exhibition 1. Team roll-up evaluations prior to Exhibition 1 is the major heading for each team evaluation. C (control), CS1 (waypoints--control/score), CS2 (scoring--control/score), S (score), CS--Control/scoring are the goals which when combined become the Complete goal rollup. Note that CS1 and CS2 are the two categories within the control/score goal and are described in the text.

Exhibition 1 results of non-participants are the results assumed for Exhibition 1 for the participating teams. This was done because the participating teams (including team 6) were not prepared to compete in Exhibition 1 but required results from Exhibition 1 such that a starting point for Bayesian updating would exist.

Team 1 roll-up evaluations prior to Exhibition 1:

	C (control)		Bins 1-10								
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0	0	0	0	0.00075	0.0025	0.0045	0.16625	0.5375	0.2885	1
Team 2	0	0	0	0	0	0	0	0.007	0.0285	0.9645	1
Team 3	0	0	0	0	0	0	0	0.007	0.021	0.972	1
Team 5	0	0	0	0	0	0	0	0.004655	0.011655	0.983655	0.999965
Team 6	0	0	0	0	0	0	0.001	0.005325	0.011325	0.982325	0.999975
Team 10	0	0	0	0	0	0	0.0025	0.00633	0.01083	0.98033	0.99999

	CS1 (waypoints--control/score)			Bins 1-10							
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0	0.018	0.09	0.156	0.232	0.268	0.18	0.056	0	0	1
Team 2	0	0	0.036	0.066	0.096	0.12	0.175	0.189	0.239	0.079	1
Team 3	0.06	0.06	0.096	0.126	0.156	0.138	0.133	0.087	0.077	0.067	1
Team 5	0	0	0.036	0.066	0.096	0.12	0.175	0.189	0.239	0.079	1
Team 6	0.01	0.01	0.07	0.1	0.13	0.082	0.112	0.172	0.232	0.082	1
Team 10	0.01	0.01	0.043	0.073	0.106	0.148	0.268	0.178	0.112	0.052	1

CS2 (scoring--control/score) **Bins 1-10**

	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.031	0.1565	0.2915	0.2815	0.1735	0.066	0	0	0	0	1
Team 2	0.018	0.0555	0.0465	0.0415	0.023	0.083	0.185	0.2625	0.182	0.103	1
Team 3	0.089	0.1305	0.1195	0.1085	0.0905	0.115	0.103	0.091	0.081	0.072	1
Team 5	0.053	0.0905	0.0835	0.1005	0.1175	0.177	0.144	0.111	0.078	0.045	1
Team 6	0.048	0.0615	0.111	0.1885	0.186	0.113	0.097	0.081	0.065	0.049	1
Team 10	0.063	0.078	0.129	0.178	0.147	0.113	0.097	0.081	0.065	0.049	1

S (score) **Bins 1-10**

	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.1674	0.2734	0.2184	0.1634	0.0853	0.0281	0.016	0.016	0.016	0.016	1
Team 2	0.1492	0.2442	0.1782	0.1122	0.0594	0.0352	0.0275	0.10508	0.06428	0.02458	0.99984
Team 3	0.081	0.275	0.209	0.143	0.077	0.011	0.011	0.09628	0.06428	0.03228	0.99984
Team 5	0.1492	0.2442	0.1804	0.1232	0.066	0.0286	0.022	0.10068	0.06208	0.02348	0.99984
Team 6	0.1602	0.2552	0.1956	0.136	0.0764	0.0366	0.043	0.0494	0.0318	0.0158	1
Team 10	0.1404	0.176	0.156	0.136	0.116	0.096	0.076	0.056	0.0318	0.0158	1

CS--Control/scoring **Bins 1-10**

	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.0217	0.11495	0.23105	0.24385	0.19105	0.1266	0.054	0.0168	0	0	1
Team 2	0.0126	0.03885	0.04335	0.04885	0.0449	0.0941	0.182	0.24045	0.1991	0.0958	1
Team 3	0.0803	0.10935	0.11245	0.11375	0.11015	0.1219	0.112	0.0898	0.0798	0.0705	1
Team 5	0.0371	0.06335	0.06925	0.09015	0.11105	0.1599	0.1533	0.1344	0.1263	0.0552	1
Team 6	0.0366	0.04605	0.0987	0.16195	0.1692	0.1037	0.1015	0.1083	0.1151	0.0589	1
Team 10	0.0471	0.0576	0.1032	0.1465	0.1347	0.1235	0.1483	0.1101	0.0791	0.0499	1

Goals / weighting **Complete goal rollup** **Bins 1-10**

Control

0.1

			1	2	3	4	5	6	7	8	9	10	check sum
Control/score													
	0.3	Team 1	0.10695	0.198525	0.200355	0.171195	0.10857	0.05509	0.02625	0.031265	0.06335	0.03845	1
		Team 2	0.0933	0.158175	0.119925	0.081975	0.04911	0.04935	0.0711	0.135883	0.101148	0.139938	0.999904
Score		Team 3	0.07269	0.197805	0.159135	0.119925	0.079245	0.04317	0.0402	0.085408	0.064608	0.137718	0.999904
	0.6	Team 5	0.10065	0.165525	0.129015	0.100965	0.072915	0.06513	0.05919	0.101194	0.076304	0.129014	0.999901
		Team 6	0.1071	0.166935	0.14697	0.130185	0.0966	0.05307	0.05635	0.062663	0.054743	0.125383	0.999998
		Team 10	0.09837	0.12288	0.12456	0.12555	0.11001	0.09465	0.09034	0.067263	0.043893	0.122483	0.999999

Team 3 roll-up evaluations prior to Exhibition 1:

C (control) — **Bins 1-10**

	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.16	0.24	0.32533	0.091655	0.024655	0.018325	0	0.028	0.084	0.028	0.999965
Team 2	0	0.005	0.015	0.019655	0.011655	0.168655	0.48	0.17862	0.04662	0.07462	0.999825
Team 3	0	0	0.32	0.24	0.165	0.0875	0.0125	0.03112	0.05712	0.08662	0.99986
Team 5	0.088	0.1691	0.2581	0.1892	0.1282	0.0272	0.0002	0	0.042	0.098	1
Team 6	0	0	0	0.056	0.141	0.2277	0.3152	0.12782	0.05382	0.07832	0.99986
Team 10	0	0	0	0.02715	0.01755	0.00795	0.1617	0.2421	0.3645	0.179	0.99995

CS1 (waypoints--control/score) — **Bins 1-10**

	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0	0	0.0399	0.1199	0.3399	0.38	0.12	0	0	0	0.9997
Team 2	0	0	0.036	0.066	0.096	0.156	0.346	0.18	0.12	0	1
Team 3	0.0798	0.1998	0.3198	0	0.01	0.04	0.03	0.1799	0.0999	0.0399	0.9991
Team 5	0.3458	0.2658	0.1858	0.106	0.063	0.033	0	0	0	0	0.9994
Team 6	0	0	0	0	0.036	0.141	0.151	0.3708	0.2208	0.0798	0.9994
Team 10	0	0	0	0	0.06	0.09	0.18	0.34	0.21	0.12	1

CS2 (scoring--control/score) — **Bins 1-10**

	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0	0	0.00665	0.04665	0.17985	0.2032	0.2732	0.14	0.09	0.06	0.99955
Team 2	0	0.005	0.02	0.045	0.065	0.08	0.075	0.3732	0.2232	0.1132	0.9996
Team 3	0.5065	0.3465	0.0965	0.005	0.02	0.015	0.01	0	0	0	0.9995
Team 5	0.4727	0.2867	0.1107	0.0515	0.0405	0.0285	0.009	0	0	0	0.9996
Team 6	0	0	0	0.0945	0.26	0.3352	0.1907	0.1012	0.018	0	0.9996
Team 10	0	0	0	0.04	0.06	0.11	0.3532	0.2765	0.1465	0.0133	0.9995

S (score)			Bins 1-10								
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0	0	0	0.03731	0.45209	0.31009	0.19978	0	0	0	0.99927
Team 2	0	0	0.014	0.153	0.226	0.293	0.115	0.108	0.059	0.032	1
Team 3	0.19959	0.57159	0.22859	0	0	0	0	0	0	0	0.99977
Team 5	0.7	0.3	0	0	0	0	0	0	0	0	1
Team 6	0	0	0.0084	0.0259	0.0981	0.3774	0.279	0.1827	0.0253	0.0032	1
Team 10	0	0	0	0.132	0.205	0.3244	0.1629	0.0949	0.0539	0.0269	1

CS--Control/scoring			Bins 1-10								
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0	0	0.016625	0.068625	0.227865	0.25624	0.22724	0.098	0.063	0.042	0.999595
Team 2	0	0.0035	0.0248	0.0513	0.0743	0.1028	0.1563	0.31524	0.19224	0.07924	0.99972
Team 3	0.37849	0.30249	0.16349	0.0035	0.017	0.0225	0.016	0.05397	0.02997	0.01197	0.99938
Team 5	0.43463	0.28043	0.13323	0.06785	0.04725	0.02985	0.0063	0	0	0	0.99954
Team 6	0	0	0	0.06615	0.1928	0.27694	0.17879	0.18208	0.07884	0.02394	0.99954
Team 10	0	0	0	0.028	0.06	0.104	0.30124	0.29555	0.16555	0.04531	0.99965

Goals/ weighting **Complete goal rollup** **Bins 1-10**

Control

0.1

Control/score			**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
	0.3	Team 1	0.016	0.024	0.037521	0.052139	0.342079	0.264759	0.18804	0.0322	0.0273	0.0154	0.999437
		Team 2	0	0.00155	0.01734	0.109156	0.159056	0.223506	0.16389	0.177234	0.097734	0.050434	0.999899
Score		Team 3	0.233301	0.433701	0.218201	0.02505	0.0216	0.0155	0.00605	0.019303	0.014703	0.012253	0.999662
	0.6	Team 5	0.559189	0.281039	0.065779	0.039275	0.026995	0.011675	0.00191	0	0.0042	0.0098	0.999862
		Team 6	0	0	0.00504	0.040985	0.1308	0.332292	0.252557	0.177026	0.044214	0.016934	0.999848
		Team 10	0	0	0	0.090315	0.142755	0.226635	0.204282	0.169815	0.118455	0.047633	0.99989

Team 5 roll-up evaluations prior to Exhibition 1:

C (control)			Bins 1-10								
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0	0	0.00175	0.00745	0.0412	0.0982	0.15445	0.20385	0.2571	0.236	1
Team 2	0	0	0	0	0.0028	0.0168	0.0308	0.055625	0.055225	0.838725	0.999975
Team 3	0.003	0.0226	0.0808	0.13155	0.18145	0.22145	0.26315	0.096	0	0	1
Team 5	0	0	0	0	0	0	0	0	0.243	0.757	1
Team 6	0	0	0	0.0145	0.033825	0.114725	0.269725	0.4264	0	0.14	0.999175
Team 10	0	0	0	0	0.0225	0.0225	0.245	0.56	0.003	0.147	1

CS1 (waypoints--control/score)				Bins 1-10							
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.6	0	0	0.005	0.009	0.013	0.017	0.111	0.235	0.01	1
Team 2	0	0	0.012	0.032	0.025	0.138	0.37	0.123	0.09	0.21	1
Team 3	0	0.01	0.052	0.096	0.08	0.136	0.18	0.224	0.147	0.075	1
Team 5	0	0	0	0.1733	0.2133	0.3133	0.24	0.06	0	0	0.9999
Team 6	0.3133	0.0333	0.0533	0	0	0.12	0.36	0.12	0	0	0.9999
Team 10	0	0	0	0	0.4798	0.1998	0.3198	0	0	0	0.9994

CS2 (scoring--control/score)				Bins 1-10							
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.709	0.0285	0.0405	0.0435	0.0605	0.0495	0.0385	0.021	0.009	0	1
Team 2	0.3	0	0.0532	0.1332	0.2312	0.103	0.078	0.033	0.018	0.05	0.9996
Team 3	0.036	0.1005	0.0965	0.141	0.188	0.211	0.141	0.0715	0.0095	0.0045	0.9995
Team 5	0.85	0.0105	0.0255	0.0405	0.0555	0.018	0	0	0	0	1
Team 6	0.85	0	0	0	0.15	0	0	0	0	0	1
Team 10	0.85	0.045	0.105	0	0	0	0	0	0	0	1

S (score)		Bins 1-10								
	1	2	3	4	5	6	7	8	9	10 check sum
Team 1	0.8284	0.0259	0.0189	0.0119	0.0181	0.0407	0.0297	0.0187	0.0077	0	1
Team 2	0.6684	0.08453	0.05553	0.02653	0.0049	0	0	0.048	0.112	0	0.99989
Team 3	0	0.0132	0.0352	0.0296	0.1252	0.286	0.2258	0.17	0.0904	0.0246	1
Team 5	0.77	0	0	0.014	0.042	0.014	0.02128	0.05328	0.08528	0	0.99984
Team 6	0.77	0	0	0.00931	0.02331	0.05859	0.05328	0.08528	0	0	0.99977
Team 10	0.77	0	0	0	0.048	0.112	0.021	0.049	0	0	1

CS--Control/scoring		Bins 1-10									
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.6763	0.01995	0.02835	0.03195	0.04505	0.03855	0.03205	0.048	0.0768	0.003	1
Team 2	0.21	0	0.04084	0.10284	0.16934	0.1135	0.1656	0.06	0.0396	0.098	0.99972
Team 3	0.0252	0.07335	0.08315	0.1275	0.1556	0.1885	0.1527	0.11725	0.05075	0.02565	0.99965
Team 5	0.595	0.00735	0.01785	0.08034	0.10284	0.10659	0.072	0.018	0	0	0.99997
Team 6	0.68899	0.00999	0.01599	0	0.105	0.036	0.108	0.036	0	0	0.99997
Team 10	0.595	0.0315	0.0735	0	0.14394	0.05994	0.09594	0	0	0	0.99982

Goals / weighting			Complete goal rollup		Bins 1-10							
Control												
	0.1											
Control/score			**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10** check sum
	0.3	Team 1	0.69993	0.021525	0.02002	0.01747	0.028495	0.045805	0.04288	0.046005	0.05337	0.0245	1
		Team 2	0.46404	0.050718	0.04557	0.04677	0.054022	0.03573	0.05276	0.052363	0.084603	0.113273	0.999848
Score		Team 3	0.00786	0.032185	0.054145	0.069165	0.139945	0.250295	0.207605	0.146775	0.069465	0.022455	0.999895
	0.6	Team 5	0.6405	0.002205	0.005355	0.032502	0.056052	0.040377	0.034368	0.037368	0.075468	0.0757	0.999895
		Team 6	0.668697	0.002997	0.004797	0.007036	0.048869	0.057427	0.091341	0.104608	0	0.014	0.999771
		Team 10	0.6405	0.00945	0.02205	0	0.074232	0.087432	0.065882	0.0854	0.0003	0.0147	0.999946

Team 6 roll-up evaluations prior to Exhibition 1:

Team 6 provided no evaluations prior to Exhibition 1

Team 10 roll-up evaluations prior to Exhibition 1:

C (control)	Bins 1-10										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0	0	0	0.0175	0.16	0.01	0.00583	0.00333	0.56333	0.24	0.99999
Team 2	0	0	0	0	0.00625	0.00925	0.01175	0.04225	0.6495	0.281	1
Team 3	0	0	0	0	0.01665	0.01665	0.01665	0.028	0.329	0.593	0.99995
Team 5	0	0	0	0	0.14625	0.00625	0.00625	0.014575	0.573325	0.253325	0.999975
Team 6	0	0	0	0	0.01665	0.01665	0.01665	0.03133	0.48733	0.43133	0.99994
Team 10	0	0	0	0	0.145	0.0075	0.01	0.005825	0.568325	0.263325	0.999975

CS1 (waypoints--control/score)				Bins 1-10							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0	0	0	0.06	0.39	0.35	0.15	0.04	0.01	0	1
Team 2	0	0	0	0	0	0	0.2199	0.4599	0.2899	0.03	0.9997
Team 3	0	0	0	0	0.3	0.3	0.0999	0.0999	0.1499	0.05	0.9997
Team 5	0	0	0	0	0.3	0.15	0.1833	0.1833	0.1833	0	0.9999
Team 6	0	0	0	0	0.3	0.3999	0.1332	0.1332	0.0333	0	0.9996
Team 10	0	0	0	0	0.3	0.072	0.2553	0.1953	0.1353	0.042	0.9999

CS2 (scoring--control/score)				Bins 1-10							
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.14	0.26	0	0.2257	0.2957	0.0657	0.0125	0	0	0	0.9996
Team 2	0	0	0.0333	0.0333	0.3733	0.08	0.1866	0.2266	0.0666	0	0.9997
Team 3	0	0	0.025	0.025	0.405	0.15495	0.14245	0.14245	0.0925	0.0125	0.99985
Team 5	0.26	0.14	0	0	0.16665	0.14985	0.14985	0.1332	0	0	0.99955
Team 6	0	0	0	0.0333	0.1833	0.23325	0.1998	0.1998	0.14985	0	0.9993
Team 10	0.35	0.05	0	0	0.16665	0.14985	0.14985	0.1332	0	0	0.99955

S (score)		Bins 1-10									
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.007	0.007	0.029	0.073	0.721	0.103	0.039	0.007	0.007	0.007	1
Team 2	0	0	0	0	0.79331	0.02331	0.02331	0.08528	0.05328	0.02128	0.99977
Team 3	0	0	0	0	0.77	0.02331	0.02331	0.05531	0.096	0.032	0.99993
Team 5	0	0	0	0.04	0.825	0.095	0.04	0	0	0	1
Team 6	0	0	0	0	0.42	0.42	0	0.05328	0.05328	0.05328	0.99984
Team 10	0	0	0.0032	0.0128	0.8624	0.032	0.0416	0.0288	0.016	0.0032	1

CS--Control/scoring			Bins 1-10								
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.098	0.182	0	0.17599	0.32399	0.15099	0.05375	0.012	0.003	0	0.99972
Team 2	0	0	0.02331	0.02331	0.26131	0.056	0.19659	0.29659	0.13359	0.009	0.9997
Team 3	0	0	0.0175	0.0175	0.3735	0.198465	0.129685	0.129685	0.10972	0.02375	0.999805
Team 5	0.182	0.098	0	0	0.206655	0.149895	0.159885	0.14823	0.05499	0	0.999655
Team 6	0	0	0	0.02331	0.21831	0.283245	0.17982	0.17982	0.114885	0	0.99939
Team 10	0.245	0.035	0	0	0.206655	0.126495	0.181485	0.15183	0.04059	0.0126	0.999655

Goals / weighting **Complete goal rollup** **Bins 1-10**

Control

0.1

Control/score			**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
	0.3	Team 1	0.0336	0.0588	0.0174	0.098347	0.545797	0.108097	0.040108	0.008133	0.061433	0.0282	0.999915
		Team 2	0	0	0.006993	0.006993	0.555004	0.031711	0.074138	0.14437	0.136995	0.043568	0.999772
Score		Team 3	0	0	0.00525	0.00525	0.575715	0.075191	0.054557	0.074892	0.123416	0.085625	0.999895
	0.6	Team 5	0.0546	0.0294	0	0.024	0.571622	0.102594	0.072591	0.045927	0.07383	0.025333	0.999894
		Team 6	0	0	0	0.006993	0.319158	0.338639	0.055611	0.089047	0.115167	0.075101	0.999715
		Team 10	0.0735	0.0105	0.00192	0.00768	0.593937	0.057899	0.080406	0.063412	0.07861	0.032033	0.999894

Exhibition 1 results of non-participants

	C (control)		Bins 1-10								
	1	2	3	4	5	6	7	8	9	10	check sum
Team 4	0	0	0	0	0.025	0	0	0	0	0.975	1
Team 8	1	0	0	0	0	0	0	0	0	0	1
Team 9	0	0.01	0	0	0.025	0	0	0	0	0.965	1
Team 11	0	0	0	0	0.025	0	0	0	0	0.975	1

	CS1 (waypoints--control/score)			Bins 1-10							
	1	2	3	4	5	6	7	8	9	10	check sum
Team 4	0.6	0.4	0	0	0	0	0	0	0	0	1
Team 8	1	0	0	0	0	0	0	0	0	0	1
Team 9	0	0	0	0	0	0	0.3	0.7	0	0	1
Team 11	0	0.3	0	0.1	0.6	0	0	0	0	0	1

	CS2 (scoring--control/score)			Bins 1-10							
	1	2	3	4	5	6	7	8	9	10	check sum
Team 4	0.85	0	0.15	0	0	0	0	0	0	0	1
Team 8	1	0	0	0	0	0	0	0	0	0	1
Team 9	0.35	0	0	0	0	0.1	0	0.4	0.15	0	1
Team 11	0.85	0	0.15	0	0	0	0	0	0	0	1

S (score)

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 4	1	0	0	0	0	0	0	0	0	0	1
Team 8	1	0	0	0	0	0	0	0	0	0	1
Team 9	1	0	0	0	0	0	0	0	0	0	1
Team 11	0.84	0	0	0	0	0.16	0	0	0	0	1

CS--Control/scoring **Bins 1-10**

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 4	0.775	0.12	0.105	0	0	0	0	0	0	0	1
Team 8	1	0	0	0	0	0	0	0	0	0	1
Team 9	0.245	0	0	0	0	0.07	0.09	0.49	0.105	0	1
Team 11	0.595	0.09	0.105	0.03	0.18	0	0	0	0	0	1

Goals / weighting			**Complete goal rollup**			**Bins 1-10**							
Control													
	0.1												
Control/score			**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
	0.3	Team 4	0.8325	0.036	0.0315	0	0.0025	0	0	0	0	0.0975	1
		Team 8	1	0	0	0	0	0	0	0	0	0	1
Score		Team 9	0.6735	0.001	0	0	0.0025	0.021	0.027	0.147	0.0315	0.0965	1
	0.6	Team 11	0.6825	0.027	0.0315	0.009	0.0565	0.096	0	0	0	0.0975	1

Appendix C: Data collected from Exhibition 2

Evaluations were performed by teams 1, 3, 5, 6 and 10 for Exhibition 2. All of the following applies to these teams. "Team roll-up evaluations prior to Exhibition 2" is the major heading for each team evaluation. Minor headings include C (control), CS1 (waypoints--control/score), CS2 (scoring--control/score), S (score), CS--Control/scoring, which are the goals which combine to become the Complete goal rollup. Note that CS1 and CS2 are the two categories within the control/score goal and are described in the text.

The subsequent major heading is "Bayesian projections for exhibition 2." Here the Bayesian-updated team evaluations for each of the evaluated teams are found. The Bayesian projections contain each the headings describe in the previous paragraph. This must be so since these projections start with the team evaluations as described above.

Exhibition 2 results (participants) is the next major heading. Just like the previous two paragraphs, the results contain all of the same headings as described before. The results are compared to the Bayesian projections as well as directly to the team evaluations for this event.

The next major heading is difference squared between actual results and Bayesian projections for Exhibition 2 (in terms of RMS error). This is where the comparison between actual results and the Bayesian-updated evaluations is made.

The next major heading is difference squared between actual results and team projections for Exhibition 2 (in terms of RMS error). This is where the comparison between actual results and the team evaluations is made.

Team 1 roll-up evaluations prior to Exhibition 2:

	C (control)		**Bins 1-10**								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0	0	0	0.00225	0.01925	0.07	0.1175	0.16833	0.26183	0.36058	0.99974
Team 2	0	0.00275	0.00525	0.008	0.0083	0.0972	0.35065	0.27705	0.20395	0.04685	1
Team 3	0	0	0	0.047	0.0846	0.1259	0.17575	0.21435	0.25295	0.09945	1
Team 5	0.0035	0.0035	0.0035	0.0755	0.1103	0.1507	0.1929	0.1718	0.1507	0.1296	0.992
Team 6	0.0005	0.0025	0.0045	0.0465	0.0772	0.1088	0.1421	0.17545	0.2518	0.1904	0.99975
Team 10	0.00075	0.00225	0.00375	0.00525	0.02295	0.10455	0.18435	0.28457	0.22267	0.16877	0.99986

	CS1 (waypoints--control/score)			**Bins 1-10**							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0	0	0	0.01	0.025	0.021	0.137	0.1828	0.2688	0.3548	0.9994
Team 2	0.036	0.066	0.162	0.192	0.24	0.154	0.104	0.034	0.012	0	1
Team 3	0	0.0636	0.0636	0.1	0.0886	0.1206	0.1916	0.1576	0.1236	0.0896	0.9988
Team 5	0.079	0.239	0.195	0.169	0.108	0.078	0.06	0.042	0.024	0.006	1
Team 6	0.058	0.1	0.148	0.19	0.232	0.112	0.04	0.04	0.04	0.04	1
Team 10	0.201	0.185	0.142	0.144	0.1	0.057	0.042	0.042	0.042	0.042	0.997

	CS2 (scoring--control/score)			**Bins 1-10**							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0.075	0.075	0.075	0.1	0.1375	0.1275	0.1175	0.1075	0.0975	0.0875	1
Team 2	0.018	0.064	0.157	0.23	0.313	0.15	0.058	0.008	0.002	0	1
Team 3	0.075	0.0856	0.0901	0.1051	0.126	0.1156	0.1096	0.1036	0.0976	0.0916	0.9998
Team 5	0.077	0.2295	0.2245	0.2195	0.0885	0.029	0.033	0.048	0.033	0.018	1
Team 6	0.0265	0.0595	0.0925	0.1305	0.177	0.161	0.1175	0.1125	0.0775	0.0455	1
Team 10	0.0995	0.106	0.118	0.1385	0.164	0.13	0.108	0.048	0.044	0.04	0.996

S (score)			Bins 1-10								
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.073	0.073	0.1182	0.1452	0.1722	0.1132	0.0862	0.073	0.073	0.073	1
Team 2	0.09808	0.24108	0.38518	0.061	0.1206	0.0522	0.0158	0.0114	0.007	0.007	0.99934
Team 3	0.105	0.098	0.091	0.084	0.077	0.0962	0.1362	0.1202	0.1042	0.0882	1
Team 5	0.0902	0.2882	0.2112	0.1342	0.0494	0.0212	0.037	0.0536	0.0562	0.0588	1
Team 6	0.1082	0.1115	0.1214	0.1313	0.1401	0.1186	0.1346	0.0582	0.0422	0.0262	0.9923
Team 10	0.1275	0.1279	0.1299	0.1303	0.1316	0.0888	0.066	0.066	0.066	0.066	1

CS--Control/scoring			Bins 1-10								
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.0525	0.0525	0.0525	0.073	0.10375	0.09555	0.12335	0.13009	0.14889	0.16769	0.99982
Team 2	0.0234	0.0646	0.1585	0.2186	0.2911	0.1512	0.0718	0.0158	0.005	0	1
Team 3	0.0525	0.079	0.08215	0.10357	0.11478	0.1171	0.1342	0.1198	0.1054	0.091	0.9995
Team 5	0.0776	0.23235	0.21565	0.20435	0.09435	0.0437	0.0411	0.0462	0.0303	0.0144	1
Team 6	0.03595	0.07165	0.10915	0.14835	0.1935	0.1463	0.09425	0.09075	0.06625	0.04385	1
Team 10	0.12995	0.1297	0.1252	0.14015	0.1448	0.1081	0.0882	0.0462	0.0434	0.0406	0.9963

Goals / weighting **Complete goal rollup** **Bins 1-10**

Control

0.1

Control/score			**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
	0.3	Team 1	0.05955	0.05955	0.08667	0.109245	0.13637	0.103585	0.100475	0.09966	0.11465	0.130165	0.99992
		Team 2	0.065868	0.164303	0.279183	0.10298	0.16052	0.0864	0.066085	0.039285	0.026095	0.008885	0.999604
Score		Team 3	0.07875	0.0825	0.079245	0.086171	0.089094	0.10544	0.139555	0.129495	0.119435	0.090165	0.99985
	0.6	Team 5	0.07775	0.242975	0.191765	0.149375	0.068975	0.0409	0.05382	0.0632	0.05788	0.05256	0.9992
		Team 6	0.075755	0.088645	0.106035	0.127935	0.14983	0.12593	0.123245	0.07969	0.070375	0.047915	0.995355
		Team 10	0.11556	0.115875	0.115875	0.12075	0.124695	0.096165	0.084495	0.081917	0.074887	0.068657	0.998876

Team 3 roll-up evaluations prior to Exhibition 2:

C (control)			Bins 1-10								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0	0	0.14375	0.2875	0.31625	0.115	0	0.01	0.03	0.0975	1
Team 2	0	0	0	0	0	0.0575	0.1725	0.17915	0.36165	0.22915	0.99995
Team 3	0.07625	0.16125	0.24875	0.29875	0.0775	0.01	0.03	0.01	0	0.0875	1
Team 5	0	0	0.0575	0.2875	0.4025	0.115	0	0.01	0.03	0.0975	1
Team 6	0	0	0	0.153238	0.226988	0.171988	0.2325	0.0875	0.03	0.0975	0.999713
Team 10	0	0	0	0	0	0.0575	0.1725	0.1825	0.375	0.2125	1

CS1 (waypoints--control/score)					Bins 1-10						
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0	0	0.06	0.3666	0.3266	0.2066	0.04	0	0	0	0.9998
Team 2	0	0	0	0.02	0.08	0.22	0.52	0.16	0	0	1
Team 3	0	0	0.14	0.4	0.22	0.14	0.08	0.02	0	0	1
Team 5	0	0	0.02	0.1866	0.4464	0.2664	0.0798	0	0	0	0.9992
Team 6	0	0	0	0.3398	0.3864	0.2064	0.0666	0	0	0	0.9992
Team 10	0	0	0	0	0	0.02	0.12	0.4998	0.2798	0.0798	0.9994

CS2 (scoring--control/score)					Bins 1-10						
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0.01	0.03	0.17	0.55995	0.20995	0.01995	0	0	0	0	0.99985
Team 2	0	0	0.03	0.13	0.155	0.295	0.215	0.145	0.03	0	1
Team 3	0.77	0.10995	0.08995	0.02995	0	0	0	0	0	0	0.99985
Team 5	0	0.01	0.11	0.54315	0.26315	0.07315	0	0	0	0	0.99945
Team 6	0	0	0.14	0.4983	0.2633	0.0683	0.03	0	0	0	0.9999
Team 10	0	0	0	0	0.01	0.095	0.4632	0.2882	0.1432	0	0.9996

S (score)			Bins 1-10								
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.0425	0.1275	0.158313	0.299563	0.215813	0.09375	0.0625	0	0	0	0.999938
Team 2	0	0	0	0	0.095813	0.333313	0.383313	0.1875	0	0	0.999938
Team 3	0.658313	0.020813	0.039563	0.125	0.09375	0.0625	0	0	0	0	0.999938
Team 5	0	0.113263	0.153263	0.527825	0.197063	0.008313	0	0	0	0	0.999725
Team 6	0	0.02125	0.17375	0.39875	0.33125	0.075	0	0	0	0	1
Team 10	0	0	0	0.0425	0.07625	0.1225	0.253613	0.303925	0.158925	0.041563	0.999275

CS--Control/scoring

	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.005	0.015	0.115	0.463275	0.268275	0.113275	0.02	0	0	0	0.999825
Team 2	0	0	0.015	0.075	0.1175	0.2575	0.3675	0.1525	0.015	0	1
Team 3	0.385	0.054975	0.114975	0.214975	0.11	0.07	0.04	0.01	0	0	0.999925
Team 5	0	0.005	0.065	0.364875	0.354775	0.169775	0.0399	0	0	0	0.999325
Team 6	0	0	0.07	0.41905	0.32485	0.13735	0.0483	0	0	0	0.99955
Team 10	0	0	0	0	0.005	0.0575	0.2916	0.394	0.2115	0.0399	0.9995

Goals / weighting **Complete goal rollup** **Bins 1-10**

Control

0.1

Control/score			**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	*check sum*
	0.3	Team 1	0.027	0.081	0.143863	0.34747	0.241595	0.101733	0.0435	0.001	0.003	0.00975	0.99991
		Team 2	0	0	0.0045	0.0225	0.092738	0.282988	0.357488	0.176165	0.040665	0.022915	0.999958
Score		Team 3	0.518113	0.045105	0.083105	0.169368	0.097	0.0595	0.015	0.004	0	0.00875	0.99994
	0.6	Team 5	0	0.069458	0.117208	0.454908	0.26492	0.06742	0.01197	0.001	0.003	0.00975	0.999633
		Team 6	0	0.01275	0.12525	0.380289	0.318904	0.103404	0.03774	0.00875	0.003	0.00975	0.999836
		Team 10	0	0	0	0.0255	0.04725	0.0965	0.256898	0.318805	0.196305	0.058158	0.999415

Team 5 roll-up evaluations prior to Exhibition 2:

C (control)	Bins 1-10										
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0	0.0055	0.027375	0.040375	0.057625	0.1355	0.1935	0.252625	0.204625	0.081	0.998125
Team 2	0	0.009275	0.06465	0.149025	0.214125	0.217275	0.163775	0.110275	0.053025	0.015525	0.99695
Team 3	0	0.002625	0.052	0.162	0.226	0.23575	0.184625	0.110625	0.026375	0	1
Team 5	0	0	0	0	0.188238	0.121988	0.055738	0.033688	0.193688	0.406188	0.999525
Team 6	0	0.034225	0.050975	0.071225	0.097475	0.155625	0.180725	0.196225	0.134225	0.078725	0.999425
Team 10	0	0.01875	0.1525	0.28875	0.115	0	0.0575	0.1725	0.09875	0.09625	1

CS1 (waypoints--control/score)	Bins 1-10										
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0	0	0.03	0.068	0.14	0.2	0.26	0.2	0.102	0	1
Team 2	0	0	0.136	0.144	0.158	0.142	0.126	0.112	0.096	0.084	0.998
Team 3	0	0.016	0.062	0.098	0.118	0.138	0.16	0.18	0.222	0	0.994
Team 5	0.4264	0.2664	0.1064	0.04	0.12	0.04	0	0	0	0	0.9992
Team 6	0	0.102	0.22	0.2	0.178	0.152	0.09	0.052	0.004	0	0.998
Team 10	0	0	0	0	0.02	0.24	0.54	0.2	0	0	1

CS2 (scoring--control/score)	Bins 1-10										
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0	0	0.0265	0.0715	0.121	0.1765	0.17	0.1685	0.208	0.058	1
Team 2	0	0.0045	0.029	0.1155	0.2345	0.2455	0.1935	0.116	0.053	0.008	0.9995
Team 3	0	0.0075	0.0305	0.0555	0.0785	0.121	0.196	0.171	0.18	0.16	1
Team 5	0.62315	0.30315	0.07315	0	0	0	0	0	0	0	0.99945
Team 6	0.098	0.174	0.175	0.166	0.14	0.117	0.106	0.0195	0.0005	0	0.996
Team 10	0.00665	0.06985	0.15985	0.3015	0.2733	0.1583	0.03	0	0	0	0.99945

S (score)		Bins 1-10								
	1	2	3	4	5	6	7	8	9	10 check sum
Team 1	0	0.0075	0.068613	0.2053	0.358925	0.212313	0.053125	0.068	0.0255	0	0.999275
Team 2	0	0.0025	0.02625	0.106	0.17875	0.249375	0.233375	0.1455	0.05825	0	1
Team 3	0	0.003125	0.026875	0.06125	0.05525	0.04925	0.04325	0.1218	0.25205	0.386425	0.999275
Team 5	0.590675	0.251925	0.125675	0.03125	0	0	0	0	0	0	0.999525
Team 6	0.061875	0.160438	0.159313	0.160063	0.160813	0.133375	0.104375	0.030125	0.019875	0.005625	0.995875
Team 10	0.0825	0.12375	0.165	0.08375	0.167063	0.188313	0.127063	0.0625	0	0	0.999938

CS--Control/scoring		Bins 1-10									
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0	0	0.02825	0.06975	0.1305	0.18825	0.215	0.18425	0.155	0.029	1
Team 2	0	0.00225	0.0825	0.12975	0.19625	0.19375	0.15975	0.114	0.0745	0.046	0.99875
Team 3	0	0.01175	0.04625	0.07675	0.09825	0.1295	0.178	0.1755	0.201	0.08	0.997
Team 5	0.524775	0.284775	0.089775	0.02	0.06	0.02	0	0	0	0	0.999325
Team 6	0.049	0.138	0.1975	0.183	0.159	0.1345	0.098	0.03575	0.00225	0	0.997
Team 10	0.003325	0.034925	0.079925	0.15075	0.14665	0.19915	0.285	0.1	0	0	0.999725

Goals / weighting **Complete goal rollup** **Bins 1-10**

Control

0.1

Control/score			**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
	0.3	Team 1	0	0.00505	0.05238	0.148143	0.260268	0.197413	0.115725	0.121338	0.082263	0.0168	0.999378
		Team 2	0	0.003103	0.046965	0.117428	0.187538	0.229478	0.204328	0.132528	0.062603	0.015353	0.99932
Score		Team 3	0	0.005663	0.0352	0.075975	0.085225	0.091975	0.097813	0.136793	0.214168	0.255855	0.998665
	0.6	Team 5	0.511838	0.236588	0.102338	0.02475	0.036824	0.018199	0.005574	0.003369	0.019369	0.040619	0.999465
		Team 6	0.051825	0.141085	0.159935	0.15806	0.153935	0.135938	0.110098	0.048423	0.026023	0.011248	0.996568
		Team 10	0.050498	0.086603	0.138228	0.12435	0.155733	0.172733	0.167488	0.08475	0.009875	0.009625	0.99988

Team 6 roll-up evaluations prior to Exhibition 2:

C (control)		Bins 1-10									
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0	0	0	0.01875	0.075	0.05625	0.15125	0.2675	0.28125	0.15	1
Team 2	0.03875	0.0775	0.11625	0.4425	0	0.0375	0.05625	0.08165	0.06165	0.0879	0.99995
Team 3	0	0	0	0	0	0.11625	0.326125	0.144875	0.271125	0.14125	0.999625
Team 5	0.27125	0.11625	0	0	0.02875	0.115	0.08625	0.075	0.0675	0.24	1
Team 6	0.038238	0.134488	0.308238	0.11625	0.0775	0	0	0.00665	0.09915	0.21915	0.999663
Team 10	0	0	0.206538	0.129038	0.204775	0.095738	0.069825	0.079088	0.152838	0.06125	0.999088

CS1 (waypoints--control/score)				Bins 1-10							
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0	0	0	0.04	0.06	0.08	0.06	0.12	0.22	0.42	1
Team 2	0	0	0.044	0.214	0.354	0.254	0.134	0	0	0	1
Team 3	0	0	0	0.06	0.3	0.38	0.22	0.04	0	0	1
Team 5	0	0.06	0.24	0.18	0.14	0.08	0.06	0.08	0.12	0.04	1
Team 6	0	0.14	0.44	0.3	0.12	0	0	0	0	0	1
Team 10	0.1066	0.4064	0.3064	0.1398	0.04	0	0	0	0	0	0.9992

CS2 (scoring--control/score)				Bins 1-10							
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0	0	0.03	0.126	0.276	0.396	0.116	0.056	0	0	1
Team 2	0	0.05665	0.06165	0.06665	0.115	0.2732	0.3132	0.1132	0	0	0.99955
Team 3	0	0.04	0.16	0.161	0.381	0.156	0.081	0.021	0	0	1
Team 5	0.035	0.015	0	0	0.06	0.25	0.27995	0.21995	0.13995	0	0.99985
Team 6	0	0.015	0.1399	0.3199	0.3699	0.125	0.03	0	0	0	0.9997
Team 10	0.35985	0.23985	0.1998	0.09995	0.07995	0.02	0	0	0	0	0.9994

S (score)			Bins 1-10								
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.0275	0.11	0.145	0.19	0.29	0.155	0.0825	0	0	0	1
Team 2	0.075813	0.203175	0.188175	0.251113	0.125	0.09375	0.0625	0	0	0	0.999525
Team 3	0.10375	0.365813	0.167063	0.050813	0	0	0.03125	0.125	0.09375	0.0625	0.999938
Team 5	0.113263	0.070763	0.028263	0.054863	0.137363	0.219863	0.008313	0.020813	0.127063	0.21875	0.999313
Team 6	0.02125	0.15375	0.21125	0.35125	0.18	0.0825	0	0	0	0	1
Team 10	0.401925	0.221925	0.063175	0	0	0	0	0	0.09375	0.21875	0.999525

CS--Control/scoring			Bins 1-10								
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0	0	0.015	0.083	0.168	0.238	0.088	0.088	0.11	0.21	1
Team 2	0	0.028325	0.052825	0.140325	0.2345	0.2636	0.2236	0.0566	0	0	0.999775
Team 3	0	0.02	0.08	0.1105	0.3405	0.268	0.1505	0.0305	0	0	1
Team 5	0.0175	0.0375	0.12	0.09	0.1	0.165	0.169975	0.149975	0.129975	0.02	0.999925
Team 6	0	0.0775	0.28995	0.30995	0.24495	0.0625	0.015	0	0	0	0.99985
Team 10	0.233225	0.323125	0.2531	0.119875	0.059975	0.01	0	0	0	0	0.9993

Goals / weighting **Complete goal rollup** **Bins 1-10**

Control

0.1

Control/score			**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
	0.3	Team 1	0.0165	0.066	0.0915	0.140775	0.2319	0.170025	0.091025	0.05315	0.061125	0.078	1
		Team 2	0.049363	0.138153	0.140378	0.237015	0.14535	0.13908	0.110205	0.025145	0.006165	0.00879	0.999643
Score		Team 3	0.06225	0.225488	0.124238	0.063638	0.10215	0.092025	0.096513	0.098638	0.083363	0.051625	0.999925
	0.6	Team 5	0.100333	0.065333	0.052958	0.059918	0.115293	0.192918	0.064605	0.06498	0.12198	0.16125	0.999565
		Team 6	0.016574	0.128949	0.244559	0.31536	0.189235	0.06825	0.0045	0.000665	0.009915	0.021915	0.999921
		Team 10	0.311123	0.230093	0.134489	0.048866	0.03847	0.012574	0.006983	0.007909	0.071534	0.137375	0.999414

Team 10 roll-up evaluations prior to Exhibition 2:

C (control)		Bins 1-10									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0	0.0125	0.0125	0.0125	0.0125	0	0.154375	0.2197	0.30345	0.2722	0.999725
Team 2	0	0	0.01	0.023625	0.047625	0.055125	0.149688	0.210938	0.231938	0.271063	1
Team 3	0	0	0	0	0	0	0.07125	0.12375	0.2725	0.5325	1
Team 5	0	0	0.010438	0.032188	0.073188	0.094563	0.126813	0.044063	0.135063	0.482813	0.999125
Team 6	0.3875	0	0	0.03875	0.145	0.098	0.067125	0.0175	0.025375	0.22075	1
Team 10	0.070375	0.100125	0.13675	0.113	0.08925	0.055	0.053	0.05625	0.065	0.26125	1

CS1 (waypoints--control/score)				Bins 1-10							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0	0	0.06	0.26	0.3	0.24	0.12	0.02	0	0	1
Team 2	0	0.018	0.028	0.044	0.06	0.06	0.18	0.24	0.29	0.078	0.998
Team 3	0	0	0	0.02	0.12	0.24	0.3	0.26	0.06	0	1
Team 5	0.24	0.18	0.144	0.128	0.108	0.108	0.068	0.024	0	0	1
Team 6	0	0	0.02	0.18	0.36	0.3	0.14	0	0	0	1
Team 10	0	0	0.04	0.102	0.176	0.218	0.26	0.14	0.042	0.022	1

CS2 (scoring--control/score)				Bins 1-10							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0.82	0.03	0.03	0.045	0.06	0.015	0	0	0	0	1
Team 2	0.3	0	0	0.011	0.032	0.133	0.194	0.23	0.086	0.0135	0.9995
Team 3	0.8	0	0.015	0.06	0.045	0.035	0.02	0.015	0.01	0	1
Team 5	0.868	0.033	0.048	0.033	0.018	0	0	0	0	0	1
Team 6	0.4	0.305	0.18	0.06	0.04	0.015	0	0	0	0	1
Team 10	0.5	0	0.0175	0.0365	0.057	0.082	0.094	0.106	0.07	0.037	1

S (score) **Bins 1-10**

	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.98125	0.01875	0	0	0	0	0	0	0	0	1
Team 2	0.009375	0.015125	0.106125	0.279625	0.123125	0.049125	0.099188	0.138313	0.179563	0	0.999563
Team 3	0.816563	0.122813	0.054063	0.00625	0	0	0	0	0	0	0.999688
Team 5	0.66875	0.053125	0.065625	0.1	0.06875	0.0375	0.00625	0	0	0	1
Team 6	0.95	0.0375	0.0125	0	0	0	0	0	0	0	1
Team 10	0.58625	0.0825	0.0125	0.00625	0	0	0.0625	0.09375	0.125	0.03125	1

CS--Control/scoring **Bins 1-10**

	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.41	0.015	0.045	0.1525	0.18	0.1275	0.06	0.01	0	0	1
Team 2	0.15	0.009	0.014	0.0275	0.046	0.0965	0.187	0.235	0.188	0.04575	0.99875
Team 3	0.4	0	0.0075	0.04	0.0825	0.1375	0.16	0.1375	0.035	0	1
Team 5	0.554	0.1065	0.096	0.0805	0.063	0.054	0.034	0.012	0	0	1
Team 6	0.2	0.1525	0.1	0.12	0.2	0.1575	0.07	0	0	0	1
Team 10	0.25	0	0.02875	0.06925	0.1165	0.15	0.177	0.123	0.056	0.0295	1

Goals / weighting **Complete goal rollup** **Bins 1-10**

Control

0.1

Control/score			**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
	0.3	Team 1	0.71175	0.017	0.01475	0.047	0.05525	0.03825	0.033438	0.02497	0.030345	0.02722	0.999973
		Team 2	0.050625	0.011775	0.068875	0.178388	0.092438	0.063938	0.130581	0.174581	0.187331	0.040831	0.999363
Score		Team 3	0.609938	0.073688	0.034688	0.01575	0.02475	0.04125	0.055125	0.053625	0.03775	0.05325	0.999813
	0.6	Team 5	0.56745	0.063825	0.069219	0.087369	0.067469	0.048156	0.026631	0.008006	0.013506	0.048281	0.999913
		Team 6	0.66875	0.06825	0.0375	0.039875	0.0745	0.05705	0.027713	0.00175	0.002538	0.022075	1
		Team 10	0.433788	0.059513	0.0298	0.035825	0.043875	0.0505	0.0959	0.098775	0.0983	0.053725	1

Bayesian projections for exhibition 2 **Bins 1-10**

Team 1

	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.853571	0.031803	0.031595	0.003857	0.016713	0.016134	0.001774	0.011504	0.004995	0.028055	1
Team 2	0.761674	0.025919	0.019344	0.001889	0.007733	0.014783	0.004915	0.051143	0.008158	0.104442	1
Team 3	0.721399	0.039403	0.031205	0.003359	0.015169	0.015721	0.003378	0.039078	0.006334	0.124952	1
Team 5	0.78438	0.025892	0.019866	0.002221	0.01096	0.018625	0.003906	0.036358	0.005875	0.091918	1
Team 6	0.805855	0.025212	0.02185	0.002765	0.01402	0.014653	0.00359	0.021737	0.004069	0.086249	1
Team 10	0.788573	0.019772	0.019729	0.002841	0.01701	0.027842	0.006132	0.024859	0.003476	0.089765	1

Team 3 **Bins 1-10**

	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.416256	0.012533	0.019287	0.003829	0.171655	0.252749	0.041426	0.038621	0.007017	0.036628	1
Team 2	0	0.001149	0.012649	0.011375	0.113263	0.302788	0.051237	0.301667	0.035647	0.170226	1
Team 3	0.934773	0.03488	0.017274	0.000283	0.001669	0.002279	0.000205	0.003566	0.000582	0.004488	1
Team 5	0.984239	0.009929	0.002288	0.000195	0.000916	0.000754	2.85E-05	0	7.3E-05	0.001577	1
Team 6	0	0	0.003659	0.004251	0.092697	0.44801	0.078579	0.299873	0.016049	0.056883	1
Team 10	0	0	0	0.009653	0.104265	0.314908	0.065504	0.296459	0.044313	0.164899	1

Team 5 **Bins 1-10**

	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.988199	0.00061	0.000558	6.96E-05	0.000776	0.002373	0.000513	0.002995	0.000744	0.003162	1
Team 2	0.961749	0.00211	0.001866	0.000274	0.00216	0.002717	0.000926	0.005003	0.001732	0.021463	1
Team 3	0.23877	0.019625	0.032499	0.005931	0.081998	0.279003	0.053404	0.205561	0.020847	0.062362	1
Team 5	0.980881	6.78E-05	0.000162	0.00014	0.001656	0.002269	0.000446	0.002638	0.001142	0.010599	1
Team 6	0.98512	8.86E-05	0.00014	2.93E-05	0.001389	0.003104	0.001139	0.007105	0	0.001886	1
Team 10	0.982954	0.000291	0.000669	0	0.002197	0.004924	0.000856	0.006042	4.55E-06	0.002062	1

Team 6					Bins 1-10						
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0.34235	0.027487	0.037142	0.008238	0.096576	0.129712	0.021974	0.112321	0.026166	0.198034	1
Team 2	0.740109	0.039141	0.041194	0.012078	0.03153	0.060776	0.011748	0.024672	0.00273	0.036023	1
Team 3	0.676627	0.048934	0.025651	0.001749	0.017133	0.042346	0.015857	0.067388	0.016521	0.087795	1
Team 5	0.814142	0.00898	0.004335	0.000738	0.010798	0.037459	0.003119	0.016685	0.00624	0.097505	1
Team 6	0.524693	0.063259	0.120151	0.018264	0.074308	0.046834	0.000603	0.002184	0.006978	0.142726	1
Team 10	0.915141	0.013026	0.010528	0.000654	0.005234	0.004166	0.000656	0.004043	0.003386	0.043166	1

Team 10					Bins 1-10						
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0.624591	0.02194	0.006391	0.00516	0.195694	0.073734	0.006313	0.00697	0.011282	0.047925	1
Team 2	0	0	0.005606	0.000801	0.434336	0.047212	0.025472	0.270053	0.054912	0.161607	1
Team 3	0	0	0.00385	0.00055	0.412129	0.1024	0.017146	0.128145	0.045251	0.290529	1
Team 5	0.720076	0.007783	0	0.000893	0.145406	0.049648	0.008107	0.027924	0.009619	0.030543	1
Team 6	0	0	0	0.000633	0.197421	0.398507	0.015102	0.131659	0.036488	0.22019	1
Team 10	0.776461	0.002226	0.000401	0.000229	0.121021	0.022444	0.007193	0.030884	0.008204	0.030937	1

Exhibition 2 results (participants):

C (control)	Bins 1-10										
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	1	0	0	0	0	0	0	0	0	0	1
Team 2	0.025	0	0	0	0	0	0	0	0	0.975	1
Team 3	0	0	0	0.8	0.14	0.025	0	0	0	0.035	1
Team 5	1	0	0	0	0	0	0	0	0	0	1
Team 6	0.025	0	0.8	0	0.14	0	0	0	0	0.035	1
Team 10	0.025	0	0	0	0	0	0	0	0	0.975	1

CS1 (waypoints--control/score)	Bins 1-10										
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	1	0	0	0	0	0	0	0	0	0	1
Team 2	0.6	0	0.3	0	0.1	0	0	0	0	0	1
Team 3	0	0	0.3	0	0.1	0.6	0	0	0	0	1
Team 5	1	0	0	0	0	0	0	0	0	0	1
Team 6	1	0	0	0	0	0	0	0	0	0	1
Team 10	0	0	0	0	0	0	0.4	0.6	0	0	1

CS2 (scoring--control/score)	Bins 1-10										
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	1	0	0	0	0	0	0	0	0	0	1
Team 2	0.85	0	0	0	0.15	0	0	0	0	0	1
Team 3	0.85	0	0	0	0.15	0	0	0	0	0	1
Team 5	1	0	0	0	0	0	0	0	0	0	1
Team 6	1	0	0	0	0	0	0	0	0	0	1
Team 10	0	0	0	0	0	0.7	0.3	0	0	0	1

S (score)	Bins 1-10										
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	1	0	0	0	0	0	0	0	0	0	1
Team 2	1	0	0	0	0	0	0	0	0	0	1
Team 3	0.84	0	0	0	0	0.16	0	0	0	0	1
Team 5	1	0	0	0	0	0	0	0	0	0	1
Team 6	1	0	0	0	0	0	0	0	0	0	1
Team 10	0.89	0	0	0	0	0.11	0	0	0	0	1

CS--Control/scoring	Bins 1-10										
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	1	0	0	0	0	0	0	0	0	0	1
Team 2	0.775	0	0.09	0	0.135	0	0	0	0	0	1
Team 3	0.595	0	0.09	0	0.135	0.18	0	0	0	0	1
Team 5	1	0	0	0	0	0	0	0	0	0	1
Team 6	1	0	0	0	0	0	0	0	0	0	1
Team 10	0	0	0	0	0	0.49	0.33	0.18	0	0	1

Goals / weighting **Complete goal rollup** **Bins 1-10**

Control

0.1

Control/score			**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
	0.3	Team 1	1	0	0	0	0	0	0	0	0	0	1
		Team 2	0.835	0	0.027	0	0.0405	0	0	0	0	0.0975	1
Score		Team 3	0.6825	0	0.027	0.08	0.0545	0.1525	0	0	0	0.0035	1
	0.6	Team 5	1	0	0	0	0	0	0	0	0	0	1
		Team 6	0.9025	0	0.08	0	0.014	0	0	0	0	0.0035	1
		Team 10	0.5365	0	0	0	0	0.213	0.099	0.054	0	0.0975	1

difference squared between actual results and Bayesian projections for exhibition 2
(far right gives rms error)

Team 1					Bins 1-10						
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**rms error**
Team 1	0.021	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.050
Team 2	0.005	0.001	0.000	0.000	0.001	0.000	0.000	0.003	0.000	0.000	0.032
Team 3	0.002	0.002	0.000	0.006	0.002	0.019	0.000	0.002	0.000	0.015	0.067
Team 5	0.046	0.001	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.008	0.076
Team 6	0.009	0.001	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.007	0.046
Team 10	0.064	0.000	0.000	0.000	0.000	0.034	0.009	0.001	0.000	0.000	0.104

Team 3					Bins 1-10						
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**rms error**
Team 1	0.341	0.000	0.000	0.000	0.029	0.064	0.002	0.001	0.000	0.001	0.210
Team 2	0.697	0.000	0.000	0.000	0.005	0.092	0.003	0.091	0.001	0.005	0.299
Team 3	0.064	0.001	0.000	0.006	0.003	0.023	0.000	0.000	0.000	0.000	0.098
Team 5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.006
Team 6	0.815	0.000	0.006	0.000	0.006	0.201	0.006	0.090	0.000	0.003	0.336
Team 10	0.288	0.000	0.000	0.000	0.011	0.010	0.001	0.059	0.002	0.005	0.194

Team 5					Bins 1-10						
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**rms error**
Team 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004
Team 2	0.016	0.000	0.001	0.000	0.001	0.000	0.000	0.000	0.000	0.006	0.049
Team 3	0.197	0.000	0.000	0.005	0.001	0.016	0.003	0.042	0.000	0.003	0.164
Team 5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.007
Team 6	0.007	0.000	0.006	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.037
Team 10	0.199	0.000	0.000	0.000	0.000	0.043	0.010	0.002	0.000	0.009	0.162

Team 6					**Bins 1-10**						
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**rms error**
Team 1	0.433	0.001	0.001	0.000	0.009	0.017	0.000	0.013	0.001	0.039	0.227
Team 2	0.009	0.002	0.000	0.000	0.000	0.004	0.000	0.001	0.000	0.004	0.044
Team 3	0.000	0.002	0.000	0.006	0.001	0.012	0.000	0.005	0.000	0.007	0.059
Team 5	0.035	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.010	0.068
Team 6	0.143	0.004	0.002	0.000	0.004	0.002	0.000	0.000	0.000	0.019	0.132
Team 10	0.143	0.000	0.000	0.000	0.000	0.044	0.010	0.002	0.000	0.003	0.142

Team 10					**Bins 1-10**						
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**rms error**
Team 1	0.141	0.000	0.000	0.000	0.038	0.005	0.000	0.000	0.000	0.002	0.137
Team 2	0.697	0.000	0.000	0.000	0.155	0.002	0.001	0.073	0.003	0.004	0.306
Team 3	0.466	0.000	0.001	0.006	0.128	0.003	0.000	0.016	0.002	0.082	0.265
Team 5	0.078	0.000	0.000	0.000	0.021	0.002	0.000	0.001	0.000	0.001	0.102
Team 6	0.815	0.000	0.006	0.000	0.034	0.159	0.000	0.017	0.001	0.047	0.329
Team 10	0.058	0.000	0.000	0.000	0.015	0.036	0.008	0.001	0.000	0.004	0.110

difference squared between actual results and team projections for exhibition2
far right gives rms error

Team 1

	Bins 1-10										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**rms error**
Team 1	0.884	0.004	0.008	0.012	0.019	0.011	0.010	0.010	0.013	0.017	0.314
Team 2	0.592	0.027	0.064	0.011	0.014	0.007	0.004	0.002	0.001	0.008	0.270
Team 3	0.365	0.007	0.003	0.000	0.001	0.002	0.019	0.017	0.014	0.008	0.209
Team 5	0.851	0.059	0.037	0.022	0.005	0.002	0.003	0.004	0.003	0.003	0.314
Team 6	0.684	0.008	0.001	0.016	0.018	0.016	0.015	0.006	0.005	0.002	0.278
Team 10	0.177	0.013	0.013	0.015	0.016	0.014	0.000	0.001	0.006	0.001	0.160

Team 3

	Bins 1-10										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**rms error**
Team 1	0.947	0.007	0.021	0.121	0.058	0.010	0.002	0.000	0.000	0.000	0.341
Team 2	0.697	0.000	0.001	0.001	0.003	0.080	0.128	0.031	0.002	0.006	0.308
Team 3	0.027	0.002	0.003	0.008	0.002	0.009	0.000	0.000	0.000	0.000	0.071
Team 5	1.000	0.005	0.014	0.207	0.070	0.005	0.000	0.000	0.000	0.000	0.361
Team 6	0.815	0.000	0.002	0.145	0.093	0.011	0.001	0.000	0.000	0.000	0.327
Team 10	0.288	0.000	0.000	0.001	0.002	0.014	0.025	0.070	0.039	0.002	0.210

Team 5

	Bins 1-10										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**rms error**
Team 1	1.000	0.000	0.003	0.022	0.068	0.039	0.013	0.015	0.007	0.000	0.342
Team 2	0.697	0.000	0.000	0.014	0.022	0.053	0.042	0.018	0.004	0.007	0.293
Team 3	0.466	0.000	0.000	0.000	0.001	0.004	0.010	0.019	0.046	0.064	0.247
Team 5	0.238	0.056	0.010	0.001	0.001	0.000	0.000	0.000	0.000	0.002	0.176
Team 6	0.724	0.020	0.006	0.025	0.020	0.018	0.012	0.002	0.001	0.000	0.288
Team 10	0.236	0.007	0.019	0.015	0.024	0.002	0.005	0.001	0.000	0.008	0.178

Team 6	Bins 1-10										
	1	2	3	4	5	6	7	8	9	10	rms error
Team 1	0.967	0.004	0.008	0.020	0.054	0.029	0.008	0.003	0.004	0.006	0.332
Team 2	0.617	0.019	0.013	0.056	0.011	0.019	0.012	0.001	0.000	0.008	0.275
Team 3	0.385	0.051	0.009	0.000	0.002	0.004	0.009	0.010	0.007	0.002	0.219
Team 5	0.809	0.004	0.003	0.004	0.013	0.037	0.004	0.004	0.015	0.026	0.303
Team 6	0.785	0.017	0.027	0.099	0.031	0.005	0.000	0.000	0.000	0.000	0.310
Team 10	0.051	0.053	0.018	0.002	0.001	0.040	0.008	0.002	0.005	0.002	0.135

Team 10	Bins 1-10										
	1	2	3	4	5	6	7	8	9	10	rms error
Team 1	0.083	0.000	0.000	0.002	0.003	0.001	0.001	0.001	0.001	0.001	0.097
Team 2	0.615	0.000	0.002	0.032	0.003	0.004	0.017	0.030	0.035	0.003	0.272
Team 3	0.005	0.005	0.000	0.004	0.001	0.012	0.003	0.003	0.001	0.002	0.062
Team 5	0.187	0.004	0.005	0.008	0.005	0.002	0.001	0.000	0.000	0.002	0.146
Team 6	0.055	0.005	0.002	0.002	0.004	0.003	0.001	0.000	0.000	0.000	0.084
Team 10	0.011	0.004	0.001	0.001	0.002	0.026	0.000	0.002	0.010	0.002	0.076

Appendix D: Competition questionnaire

The questionnaire used in the Competition is presented below. The questionnaire shows the questions used to project a robot's potential prior to the Competition and to identify a robot's performance in the Competition. Minor changes were made between the questionnaire used for the Exhibitions and that used for the Competition (Appendix A). These changes are discussed in the text.

Questionnaire to be filled out by students of 6.186 who are doing the special project also.

Fill out one for your team and then one for each of the participating teams. The participating teams are 1, 2, 3, 5, 6, 10
Team members should split the work so only 2 or 3 questionnaires maximum will be filled by any one person.

Return to a staff member or Bill Hardman prior to the competition of 1/29/03.

Team filling out the questionnaire _______ 1 2 3 5 6 10
(circle team number for which questionnaire filled out)

1	**robot requires more than 60 seconds calibration**				**average performance (requires between 1 sec< time< 60 sec)**					**requires no calibration**	
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	
select range											
select peak											

2	*robot begins computing or starts* prior to 10 second required delay	*average* performance (moves starting between 30 and 10 sec)	*begins moving at exactly* 10 seconds

1 2 3 4 5 6 7 8 9 10

select range

select peak

3	robot moves after 4 min allowed time limit	average performance (stops with time to spare)	robot stops exactly at 4 min allowed time limit

1 2 3 4 5 6 7 8 9 10

select range

select peak

4	does not arrive (i.e, within 4") at any specified waypoints	average performance	arrives (i.e., within 4") at all specified waypoints

1 2 3 4 5 6 7 8 9 10

select range

select peak

5	does not perform waypoint signal after arriving (i.e, within 4") at any specified waypoints	average performance	performs waypoint signal every time within 4" of a waypoint

1 2 3 4 5 6 7 8 9 10

select range

select peak

6	retrieves and places in yellow scoring locations far less than average # of targets	average performance	retrieves and places in yellow areas greater than average # of targets

1 2 3 4 5 6 7 8 9 10

select range

select peak

7	robot unable to get to yellow scoring area	average performance	robot always arrives at yellow scoring area

1 2 3 4 5 6 7 8 9 10

select range

select peak

8	robot unable to detach target to deposit in yellow scoring area	average performance	robot always detaches target for deposit in scoring area

1 2 3 4 5 6 7 8 9 10

select range

select peak

9	retrieved and placed in home scoring location far less than average # of targets	average performance	retrieved and placed in home scoring area greater than average # of targets

1 2 3 4 5 6 7 8 9 10

select range

select peak

10	when robot fails to grasp target not smart enough to try again				average performance				when robot fails to grasp target is smart enough to try again			
	1	2	3	4	5	6	7	8	9	10		

select range
select peak

11	collides with all large objects placed in its path				average performance				collides with no objects placed in its path			
	1	2	3	4	5	6	7	8	9	10		

select range
select peak

12	robot unreliable in detecting scoring areas				average performance				robot 100% reliable in detecting scoring areas			
	1	2	3	4	5	6	7	8	9	10		

select range
select peak

13	robot unreliable in detecting the targets				average performance				robot 100% reliable in detecting the targets			
	1	2	3	4	5	6	7	8	9	10		

select range
select peak

14 robot has long processing time in detecting scoring areas and/or targets | average performance | robot has short processing time in detecting scoring areas and/or targets

1	2	3	4	5	6	7	8	9	10

select range
select peak

15 small percentage of playing field explored | average performance | large percentage of playing field explored

1	2	3	4	5	6	7	8	9	10

select range
select peak

16 robot does not arrive home in 4 minutes allowed | average performance | robot arrives home with time to spare

1	2	3	4	5	6	7	8	9	10

select range
select peak

17 robot cannot operate reliably on battery power | average performance | robot operates reliably on battery power

1	2	3	4	5	6	7	8	9	10

select range
select peak

18	robot team 1 is last in competition				average performance				robot team 1 wins competition		
	1	2	3	4	5	6	7	8	9	10	
select range											
select peak											

19	robot team 2 is last in competition				average performance				robot team 2 wins competition		
	1	2	3	4	5	6	7	8	9	10	
select range											
select peak											

20	robot team 3 is last in competition				average performance				robot team 3 wins competition		
	1	2	3	4	5	6	7	8	9	10	
select range											
select peak											

21	robot team 5 is last in competition				average performance				robot team 5 wins competition		
	1	2	3	4	5	6	7	8	9	10	
select range											
select peak											

22 robot team 6 is last in competition | average performance | robot team 6 wins competition

1	2	3	4	5	6	7	8	9	10

select range

select peak

23 robot team 10 is last in competition | average performance | robot team 10 wins competition

1	2	3	4	5	6	7	8	9	10

select range

select peak

Appendix E: Data collected from Competition

Evaluations were performed by teams 1, 3, 5, 6 and 10 for the Competition. All of the following applies to these teams. "Team roll-up evaluations prior to the Competition" is the major heading for each team evaluation. Minor headings include C (control), CS1 (waypoints--control/score), CS2 (scoring--control/score), S (score), CS--Control/scoring, which are the goals which combine to become the Complete goal rollup. Note that CS1 and CS2 are the two categories within the control/score goal and are described in the text.

Note also, that a minor heading of "final score" is found between S (score) and CS--Control/scoring. This was data not used in the project but was asked for in the Competition questionnaire as the teams one question rollup, (i.e., the evaluating team's overall opinion) of the teams prior to the Competition.

The subsequent major heading is "Bayesian projections for the Competition." Here the Bayesian-updated team evaluations for each of the evaluated teams are found. The Bayesian projections contain each the headings describe in the previous paragraph. This must be so since these projections start with the team evaluations as described above.

Competition results (participants) is the next major heading. Just like the previous two paragraphs, the results contain all of the same headings as described before. The results are compared to the Bayesian projections as well as directly to the team evaluations for this event.

The next major heading is difference squared between actual results and Bayesian projections for the Competition (in terms of RMS error). This is where the comparison between actual results and the Bayesian-updated evaluations is made.

The next major heading is difference squared between actual results and team projections for the Competition (in terms of RMS error). This is where the comparison between actual results and the team evaluations is made.

Team 1 roll-up evaluations prior to Competition:

	C (control)		Bins 1-10								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0	0	0	0	0	0	0.16	0.48999	0.18799	0.16199	0.99997
Team 2	0	0	0.188	0.564	0.188	0	0.006	0.02199	0.01599	0.01599	0.99997
Team 3	0.014	0.014	0.03	0.0807	0.1335	0.1887	0.2358	0.1685	0.1009	0.0336	0.9997
Team 5	0	0	0	0	0	0.188	0.564	0.19799	0.02799	0.02199	0.99997
Team 6	0.014	0.014	0.018	0.0367	0.0732	0.1014	0.1379	0.1657	0.2415	0.1893	0.9917
Team 10	0.014	0.014	0.018	0.0349	0.072	0.1008	0.1379	0.1687	0.2433	0.1884	0.992

	CS1 (waypoints--control/score)				Bins 1-10						
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0	0	0	0.14	0.42	0.14	0	0.06	0.18	0.06	1
Team 2	0.1599	0.0999	0.0399	0.14	0.42	0.14	0	0	0	0	0.9997
Team 3	0.03	0.03	0.03	0.27	0.21	0.157	0.107	0.057	0.067	0.042	1
Team 5	0	0.06	0.3	0.44	0.18	0.02	0	0	0	0	1
Team 6	0.03	0.03	0.03	0.27	0.21	0.152	0.101	0.05	0.059	0.067	0.999
Team 10	0.03	0.03	0.03	0.27	0.21	0.157	0.107	0.057	0.067	0.042	1

	CS2 (scoring--control/score)				Bins 1-10						
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0.3731	0.2331	0.0931	0.06	0.18	0.06	0	0	0	0	0.9993
Team 2	0.02665	0.01665	0.03665	0.09	0.17	0.44	0.2	0.02	0	0	0.99995
Team 3	0.09	0.09	0.1175	0.1115	0.1075	0.1035	0.0995	0.0965	0.0925	0.0905	0.999
Team 5	0.01	0.01	0.02	0.04	0.11	0.28	0.18	0.25	0.09	0.01	1
Team 6	0.1	0.1	0.1	0.1	0.1	0.105	0.1	0.1	0.1	0.095	1
Team 10	0.1	0.1	0.1	0.1	0.105	0.1	0.1	0.1	0.1	0.095	1

S (score)				Bins 1-10							
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.43706	0.27306	0.10906	0	0	0.03	0.09	0.03399	0.00999	0.01599	0.99915
Team 2	0.144	0.432	0.144	0.03	0.104	0.092	0.08	0.038	0.006	0	1.07
Team 3	0.074	0.08	0.0855	0.0915	0.0981	0.1101	0.0951	0.1116	0.1218	0.1308	0.9985
Team 5	0	0	0	0.034	0.132	0.254	0.426	0.148	0.006	0	1
Team 6	0.083	0.093	0.104	0.094	0.0846	0.0881	0.1041	0.1221	0.1338	0.0933	1
Team 10	0.084	0.104	0.097	0.09	0.0826	0.0816	0.0951	0.1116	0.1218	0.1308	0.9985

final score		(student single		question guess)			Bins 1-10				
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0	0	0	0	0	0.06	0.11	0.2265	0.2765	0.3265	0.9995
Team 2	0	0	0.015	0.065	0.12	0.1	0.08	0.16	0.34	0.12	1
Team 3	0	0	0	0.1	0.355	0.205	0.16	0.11	0.06	0.01	1
Team 5	0	0	0	0	0.1	0.36	0.285	0.135	0.085	0.035	1
Team 6	0	0	0.1375	0.2075	0.3875	0.1675	0.0475	0.0325	0.0125	0.0025	0.995
Team 10	0.2665	0.2165	0.1165	0.05	0.05	0.1	0.05	0.05	0.05	0.05	0.9995

CS--Control/scoring					Bins 1-10						
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.26117	0.16317	0.06517	0.084	0.252	0.084	0	0.018	0.054	0.018	0.99951
Team 2	0.066625	0.041625	0.037625	0.105	0.245	0.35	0.14	0.014	0	0	0.999875
Team 3	0.072	0.072	0.09125	0.15905	0.13825	0.11955	0.10175	0.08465	0.08485	0.07595	0.9993
Team 5	0.007	0.025	0.104	0.16	0.131	0.202	0.126	0.175	0.063	0.007	1
Team 6	0.079	0.079	0.079	0.151	0.133	0.1191	0.1003	0.085	0.0877	0.0866	0.9997
Team 10	0.079	0.079	0.079	0.151	0.1365	0.1171	0.1021	0.0871	0.0901	0.0791	1

Goals / weighting **Complete goal rollup of** **Bins 1-10**

Control

0.1

Control/score			**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
	0.3	Team 1	0.340587	0.212787	0.084987	0.0252	0.0756	0.0432	0.07	0.074793	0.040993	0.031193	0.99934
		Team 2	0.106388	0.271688	0.116488	0.1059	0.1547	0.1602	0.0906	0.029199	0.005199	0.001599	1.04196
Score		Team 3	0.0674	0.071	0.081675	0.110685	0.113685	0.120795	0.111165	0.109205	0.108625	0.104625	0.99886
	0.6	Team 5	0.0021	0.0075	0.0312	0.0684	0.1185	0.2318	0.3498	0.161099	0.025299	0.004299	0.999997
		Team 6	0.0749	0.0809	0.0879	0.10537	0.09798	0.09873	0.10634	0.11533	0.13074	0.10089	0.99908
		Team 10	0.0755	0.0875	0.0837	0.10279	0.09771	0.09417	0.10148	0.10996	0.12444	0.12105	0.9983

Team 3 roll-up evaluations prior to Competition:

C (control)				Bins 1-10							
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0	0.003	0.012	0.12099	0.19199	0.25999	0.176	0.096	0.042	0.098	0.99997
Team 2	0	0	0	0.062	0.151	0.237	0.311	0.099	0.042	0.098	1
Team 3	0	0	0	0	0	0.08	0.32	0.25599	0.22099	0.12299	0.99997
Team 5	0.00999	0.02799	0.11799	0.296	0.216	0.136	0.056	0	0.042	0.098	0.99997
Team 6	0	0	0	0.08	0.326	0.26199	0.17599	0.04399	0.084	0.028	0.99997
Team 10	0	0	0	0	0.08	0.32	0.249	0.19362	0.06762	0.08962	0.99986

CS1 (waypoints--control/score)				Bins 1-10							
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0	0	0.12	0.36	0.2	0.21	0.1	0.01	0	0	1
Team 2	0	0.03	0.07	0	0.06	0.09	0.4398	0.2298	0.0798	0	0.9994
Team 3	0	0	0	0.072	0.132	0.222	0.2653	0.1953	0.1133	0	0.9999
Team 5	0.18	0.432	0.07	0.09	0.113	0.073	0.036	0.006	0	0	1
Team 6	0	0	0	0.06	0.4199	0.3399	0.1799	0	0	0	0.9997
Team 10	0	0	0	0	0.0798	0.2798	0.5598	0.08	0	0	0.9994

CS2 (scoring--control/score)				Bins 1-10							
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0	0.04	0.196	0.2943	0.2093	0.1143	0.116	0.03	0	0	0.9999
Team 2	0.03	0.3332	0.2765	0.17315	0.0499	0.0566	0.07995	0	0	0	0.9993
Team 3	0	0.018	0.1929	0.2399	0.3349	0.13	0.022	0.03865	0.01665	0.00665	0.99965
Team 5	0.2233	0.5283	0.0733	0.045	0.055	0.06	0.015	0	0	0	0.9999
Team 6	0	0	0	0.48305	0.37305	0.14305	0	0	0	0	0.99915
Team 10	0	0	0	0	0	0.31985	0.41985	0.19985	0.06	0	0.99955

S (score)				Bins 1-10							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0	0.084	0.41845	0.26345	0.14345	0.069	0.021	0	0	0	0.99935
Team 2	0	0	0.1384	0.4657	0.1808	0.0445	0.0736	0.0567	0.0366	0.0036	0.9999
Team 3	0	0	0.0084	0.0354	0.0524	0.0554	0.0394	0.208	0.441	0.16	1
Team 5	0.18	0.523	0.2331	0.0341	0.0151	0.0111	0.0036	0	0	0	1
Team 6	0	0	0.34645	0.26975	0.12275	0.0693	0.141	0.05	0	0	0.99925
Team 10	0	0	0	0	0.065	0.314	0.304	0.26	0.051	0.006	1

final score		(student single		question guess)			Bins 1-10				
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0.1	0.15	0.4665	0.2165	0.0665	0	0	0	0	0	0.9995
Team 2	0	0	0	0	0	0.1	0.2	0.4	0.2	0.1	1
Team 3	0	0	0	0.035	0.085	0.135	0.185	0.3265	0.1665	0.0665	0.9995
Team 5	0.1	0.2	0.4	0.2	0.1	0	0	0	0	0	1
Team 6	0	0	0.11	0.31	0.31	0.21	0.06	0	0	0	1
Team 10	0	0	0	0	0	0.035	0.185	0.285	0.385	0.11	1

CS--Control/scoring					Bins 1-10						
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0	0.028	0.1732	0.31401	0.20651	0.14301	0.1112	0.024	0	0	0.99993
Team 2	0.021	0.24224	0.21455	0.121205	0.05293	0.06662	0.187905	0.06894	0.02394	0	0.99933
Team 3	0	0.0126	0.13503	0.18953	0.27403	0.1576	0.09499	0.085645	0.045645	0.004655	0.999725
Team 5	0.21031	0.49941	0.07231	0.0585	0.0724	0.0639	0.0213	0.0018	0	0	0.99993
Team 6	0	0	0	0.356135	0.387105	0.202105	0.05397	0	0	0	0.999315
Team 10	0	0	0	0	0.02394	0.307835	0.461835	0.163895	0.042	0	0.999505

Goals / weighting **Complete goal rollup of** **Bins 1-10**

Control

0.1

Control/score			**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
	0.3	Team 1	0	0.0591	0.30423	0.264372	0.167222	0.110302	0.06356	0.0168	0.0042	0.0098	0.999586
		Team 2	0.0063	0.072672	0.147405	0.321982	0.139459	0.070386	0.131632	0.064602	0.033342	0.01196	0.999739
Score		Team 3	0	0.00378	0.045549	0.078099	0.113649	0.08852	0.084137	0.176093	0.300393	0.109696	0.999915
	0.6	Team 5	0.172092	0.466422	0.173352	0.06761	0.05238	0.03943	0.01415	0.00054	0.0042	0.0098	0.999976
		Team 6	0	0	0.20787	0.276691	0.222382	0.128411	0.11839	0.034399	0.0084	0.0028	0.999342
		Team 10	0	0	0	0	0.054182	0.312751	0.345851	0.224531	0.049962	0.012562	0.999838

Team 5 roll-up evaluations prior to Competition:

	C (control)				Bins 1-10						
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0	0	0	0	0	0.136	0.362	0.27598	0.19398	0.03198	0.99994
Team 2	0	0	0.02	0.053	0.0354	0.1898	0.5162	0.1786	0.007	0	1
Team 3	0	0	0	0	0.24	0.56	0	0.02261	0.06561	0.11161	0.99983
Team 5	0	0	0.16	0.48	0.16	0	0	0.012	0.036	0.152	1
Team 6	0	0	0	0	0	0	0	0	0.611	0.389	1
Team 10	0	0	0	0	0	0	0	0	0.3	0.7	1

	CS1 (waypoints--control/score)				Bins 1-10						
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0	0	0	0	0	0	0.06	0.5131	0.2931	0.1331	0.9993
Team 2	0	0	0.02	0.06	0.02	0.072	0.162	0.312	0.222	0.132	1
Team 3	0	0	0	0.06	0.24	0.2	0.2199	0.1199	0.1599	0	0.9997
Team 5	0	0.18	0.56	0.24	0.02	0	0	0	0	0	1
Team 6	0	0	0.036	0.066	0.096	0.086	0.066	0.3598	0.2098	0.0798	0.9994
Team 10	0	0	0	0	0	0	0	0.1333	0.6033	0.2633	0.9999

	CS2 (scoring--control/score)				Bins 1-10						
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0	0	0	0	0.02	0.2449	0.1999	0.3181	0.1632	0.0532	0.9993
Team 2	0.04	0.166	0.1385	0.128	0.103	0.143	0.1445	0.0795	0.0445	0.01	0.997
Team 3	0.21	0.09	0.0532	0.1432	0.33315	0.07995	0.03995	0	0.015	0.035	0.99945
Team 5	0.2132	0.2232	0.3532	0.18	0.03	0	0	0	0	0	0.9996
Team 6	0.046	0.051	0.056	0.061	0.101	0.1	0.105	0.1	0.25	0.13	1
Team 10	0	0	0	0	0	0	0.03	0.3	0.39	0.28	1

	S (score)				**Bins 1-10**						
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0.13931	0.41331	0.16731	0	0	0	0	0.05399	0.15999	0.06599	0.9999
Team 2	0.34645	0.22245	0.10445	0.006	0	0.018	0.0555	0.06981	0.10881	0.06781	0.99928
Team 3	0	0	0	0.36645	0.31376	0.12976	0.00931	0.00399	0.11499	0.06099	0.99925
Team 5	0.15	0.494	0.282	0.044	0.006	0.018	0.006	0	0	0	1
Team 6	0.476	0.244	0	0.018	0.045	0.07	0.065	0.04	0.021	0.021	1
Team 10	0	0	0	0	0	0	0.01	0.496	0.34	0.154	1

	final score		**(student**	**single**	**question**	**guess)**		**Bins 1-10**			
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0	0	0	0	0	0.05	0.25	0.25	0.25	0.2	1
Team 2	0	0.2665	0.1665	0.1265	0.185	0.135	0.085	0.035	0	0	0.9995
Team 3	0	0	0.06	0.11	0.26	0.26	0.26	0.05	0	0	1
Team 5	0.56	0.185	0.135	0.085	0.035	0	0	0	0	0	1
Team 6	0	0	0	0.06	0.185	0.135	0.085	0.135	0.3	0.1	1
Team 10	0	0	0	0	0	0.16	0.41	0.26	0.11	0.06	1

CS--Control/scoring											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0	0	0	0	0.014	0.17143	0.15793	0.3766	0.20217	0.07717	0.9993
Team 2	0.028	0.1162	0.10295	0.1076	0.0781	0.1217	0.14975	0.14925	0.09775	0.0466	0.9979
Team 3	0.147	0.063	0.03724	0.11824	0.305205	0.115965	0.093935	0.03597	0.05847	0.0245	0.999525
Team 5	0.14924	0.21024	0.41524	0.198	0.027	0	0	0	0	0	0.99972
Team 6	0.0322	0.0357	0.05	0.0625	0.0995	0.0958	0.0933	0.17794	0.23794	0.11494	0.99982
Team 10	0	0	0	0	0	0	0.021	0.24999	0.45399	0.27499	0.99997

Goals / weighting **Complete goal rollup Bins 1-10**

Control

0.1

			1	2	3	4	5	6	7	8	9	10	check sum
Control/score													
	0.3	Team 1	0.083586	0.247986	0.100386	0	0.0042	0.065029	0.083579	0.172972	0.176043	0.065943	0.999724
		Team 2	0.21627	0.16833	0.095555	0.04118	0.02697	0.06629	0.129845	0.104521	0.095311	0.054666	0.998938
Score		Team 3	0.0441	0.0189	0.011172	0.255342	0.303818	0.168646	0.033767	0.015446	0.093096	0.055105	0.999391
	0.6	Team 5	0.134772	0.359472	0.309772	0.1338	0.0277	0.0108	0.0036	0.0012	0.0036	0.0152	0.999916
		Team 6	0.29526	0.15711	0.015	0.02955	0.05685	0.07074	0.06699	0.077382	0.145082	0.085982	0.999946
		Team 10	0	0	0	0	0	0	0.0123	0.372597	0.370197	0.244897	0.999991

Team 6 roll-up evaluations prior to Competition:

	C (control)			Bins 1-10							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0	0	0	0	0	0	0.4264	0.2724	0.2224	0.078	0.9992
Team 2	0.01599	0.00999	0.00399	0	0	0	0	0.03199	0.09399	0.84399	0.99994
Team 3	0	0	0	0	0	0	0	0.44901	0.33201	0.21801	0.99903
Team 5	0	0.088	0.168	0.256	0.176	0.096	0.016	0.012	0.036	0.152	1
Team 6	0	0	0	0	0	0	0	0	0	1	1
Team 10	0	0	0	0	0	0	0	0.16	0.48	0.36	1

	CS1 (waypoints--control/score)			Bins 1-10							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0	0.3	0.3399	0.0999	0.1599	0.04	0.03	0.02	0.01	0	0.9997
Team 2	0	0	0	0.09	0.21	0	0	0.02	0.24	0.44	1
Team 3	0	0	0	0.06	0.24	0.3399	0.2332	0.0732	0.0533	0	0.9996
Team 5	0.036	0.146	0.366	0.286	0.166	0	0	0	0	0	1
Team 6	0	0.006	0.03	0.054	0.088	0.1	0.132	0.284	0.186	0.12	1
Team 10	0	0	0	0	0	0	0	0.12	0.45	0.43	1

	CS2 (scoring--control/score)			Bins 1-10							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0	0	0.02665	0.01665	0.00665	0.0665	0.3564	0.4564	0.0699	0	0.99915
Team 2	0.1599	0.3131	0.1731	0.0732	0.06	0.065	0.105	0	0.015	0.035	0.9993
Team 3	0	0.0532	0.1332	0.2531	0.0999	0.1599	0.0133	0.12325	0.13325	0.02995	0.99905
Team 5	0.436	0.1	0.2045	0.1415	0.0775	0.0225	0.0135	0.0045	0	0	1
Team 6	0.005	0.023	0.042	0.059	0.151	0.172	0.193	0.052	0.003	0.3	1
Team 10	0.7	0	0.024	0.044	0.064	0.044	0.024	0.02	0.06	0.02	1

S (score) **Bins 1-10**

	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0	0.03731	0.02331	0.00931	0.26	0.195	0.175	0.2	0.07	0.03	0.99993
Team 2	0.15	0.48731	0.17331	0.00931	0	0.006	0.018	0.02595	0.04995	0.07995	0.99978
Team 3	0.06261	0.05661	0.05061	0	0	0.34645	0.24645	0.17645	0.039	0.021	0.99918
Team 5	0.1577	0.4947	0.2624	0.0454	0.0144	0.0194	0.006	0	0	0	1
Team 6	0.72	0	0.012	0.022	0.032	0.022	0.012	0	0	0.18	1
Team 10	0.1	0.007	0.028	0.021	0.014	0	0	0.13	0.39	0.31	1

final score **(student single question guess)** **Bins 1-10**

	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.03	0.1925	0.3675	0.27575	0.10075	0.03325	0	0	0	0	0.99975
Team 2	0	0	0.08	0.23	0.36325	0.21325	0.11325	0	0	0	0.99975
Team 3	0	0	0.16325	0.16325	0.26325	0.28	0.13	0	0	0	0.99975
Team 5	0.05	0.25	0.43325	0.23325	0.03325	0	0	0	0	0	0.99975
Team 6	0	0	0.055	0.23	0.255	0.13	0.0525	0.2025	0.075	0	1
Team 10	0	0	0	0	0	0.05	0.2	0.225	0.4	0.125	1

CS--Control/scoring **Bins 1-10**

	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0	0.09	0.120625	0.041625	0.052625	0.05855	0.25848	0.32548	0.05193	0	0.999315
Team 2	0.11193	0.21917	0.12117	0.07824	0.105	0.0455	0.0735	0.006	0.0825	0.1565	0.99951
Team 3	0	0.03724	0.09324	0.19517	0.14193	0.2139	0.07927	0.108235	0.109265	0.020965	0.999215
Team 5	0.316	0.1138	0.25295	0.18485	0.10405	0.01575	0.00945	0.00315	0	0	1
Team 6	0.0035	0.0179	0.0384	0.0575	0.1321	0.1504	0.1747	0.1216	0.0579	0.246	1
Team 10	0.49	0	0.0168	0.0308	0.0448	0.0308	0.0168	0.05	0.177	0.143	1

Goals / weighting **Complete goal rollup** **Bins 1-10**

Control

0.1

Control/score			**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
	0.3	Team 1	0	0.049386	0.050174	0.018074	0.171788	0.134565	0.225184	0.244884	0.079819	0.0258	0.999673
		Team 2	0.125178	0.359136	0.140736	0.029058	0.0315	0.01725	0.03285	0.020569	0.064119	0.179319	0.999715
Score		Team 3	0.037566	0.045138	0.058338	0.058551	0.042579	0.27204	0.171651	0.183242	0.089381	0.040691	0.999176
	0.6	Team 5	0.18942	0.33976	0.250125	0.108295	0.057455	0.025965	0.008035	0.002145	0.0036	0.0152	1
		Team 6	0.43305	0.00537	0.01872	0.03045	0.05883	0.05832	0.05961	0.03648	0.01737	0.2818	1
		Team 10	0.207	0.0042	0.02184	0.02184	0.02184	0.00924	0.00504	0.109	0.3351	0.2649	1

Team 10 roll-up evaluations prior to Competition:

C (control)				Bins 1-10						
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10** check sum
Team 1	0	0	0	0	0.4124	0.2754	0.28839	0.01299	0.00999	0	0.99917
Team 2	0	0	0	0.015	0.475	0.255	0.175	0.08	0	0	1
Team 3	0	0	0.0027	0.0039	0.35509	0.21629	0.21509	0.2039	0.0027	0	0.99967
Team 5	0	0	0.0966	0.1802	0.4062	0.1922	0.1119	0.0096	0.0033	0	1
Team 6	0	0	0.006	0.22012	0.54012	0.22012	0.0135	0	0	0	0.99986
Team 10	0.014	0.094	0.334	0.254	0.204	0.02	0.023	0.026	0.017	0.014	1

CS1 (waypoints--control/score)				Bins 1-10							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0	0	0.08	0.12	0.3598	0.2398	0.1998	0	0	0	0.9994
Team 2	0	0	0	0	0	0	0.05	0.57	0.31	0.07	1
Team 3	0	0	0.0999	0.0999	0.0999	0.15	0.15	0.1833	0.1833	0.0333	0.9996
Team 5	0	0	0	0.108	0.388	0.288	0.188	0.028	0	0	1
Team 6	0	0	0.054	0.178	0.206	0.228	0.25	0.084	0	0	1
Team 10	0.06	0.06	0.06	0.06	0.06	0.072	0.157	0.167	0.157	0.147	1

CS2 (scoring--control/score)				Bins 1-10							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
Team 1	0.025	0.065	0.21	0.31985	0.22985	0.14985	0	0	0	0	0.99955
Team 2	0	0.022	0.0565	0.091	0.1255	0.0745	0.1782	0.2332	0.1882	0.03	0.9991
Team 3	0	0	0.105	0.29485	0.33485	0.21485	0	0.01665	0.01665	0.01665	0.9995
Team 5	0	0.2025	0.288	0.273	0.0965	0.0965	0.033	0.01	0	0	0.9995
Team 6	0	0	0.0999	0.1599	0.2499	0.08	0.1432	0.1332	0.1332	0	0.9993
Team 10	0.53	0.048	0.063	0.078	0.063	0.048	0.03	0.03665	0.04665	0.05665	0.99995

	S (score)				Bins 1-10						
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.65	0	0.05661	0.0666	0.0816	0.06999	0.045	0.03	0	0	0.9998
Team 2	0.137	0.137	0.137	0.137	0.143	0.025	0.0463	0.0703	0.1303	0.037	0.9999
Team 3	0.02331	0.02331	0.02801	0.3409	0.3521	0.0563	0.0591	0.0629	0.0357	0.018	0.99963
Team 5	0.16731	0.44661	0.17861	0.0393	0.0435	0.0435	0.0435	0.0375	0	0	0.99983
Team 6	0	0	0.06021	0.40321	0.41521	0.0456	0.0576	0.018	0	0	0.99983
Team 10	0.107	0.007	0.007	0.007	0.0133	0.0313	0.0496	0.0616	0.0421	0.6741	1

	final score	(student	single	question	guess)		Bins 1-10				
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.066667	0.2	0.106667	0.073333	0.106667	0.073333	0.04	0.177667	0.111	0.044333	0.999667
Team 2	0	0	0.066667	0.333333	0.4	0.2	0	0	0	0	1
Team 3	0	0	0	0.1	0.233333	0.273333	0.173333	0.106667	0.073333	0.04	1
Team 5	0.033333	0.153333	0.163333	0.223333	0.17	0.183333	0.063333	0.01	0	0	1
Team 6	0.066667	0.13	0.21	0.223333	0.17	0.123333	0.043333	0.03	0	0	0.996667
Team 10	0.04	0.073333	0.106667	0.083333	0.073333	0.083333	0.14	0.196667	0.163333	0.04	1

CS--Control/scoring				Bins 1-10							
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.0175	0.0455	0.171	0.259895	0.268835	0.176835	0.05994	0	0	0	0.999505
Team 2	0	0.0154	0.03955	0.0637	0.08785	0.05215	0.13974	0.33424	0.22474	0.042	0.99937
Team 3	0	0	0.10347	0.236365	0.264365	0.195395	0.045	0.066645	0.066645	0.021645	0.99953
Team 5	0	0.14175	0.2016	0.2235	0.18395	0.15395	0.0795	0.0154	0	0	0.99965
Team 6	0	0	0.08613	0.16533	0.23673	0.1244	0.17524	0.11844	0.09324	0	0.99951
Team 10	0.389	0.0516	0.0621	0.0726	0.0621	0.0552	0.0681	0.075755	0.079755	0.083755	0.999965

Goals / weighting **Complete goal rollup** **Bins 1-10**

Control

0.1

Control/score			**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
	0.3	Team 1	0.39525	0.01365	0.085266	0.117929	0.170851	0.122585	0.073821	0.019299	0.000999	0	0.999649
		Team 2	0.0822	0.08682	0.094065	0.10281	0.159655	0.056145	0.087202	0.150452	0.145602	0.0348	0.999751
Score		Team 3	0.013986	0.013986	0.048117	0.27584	0.326079	0.114028	0.070469	0.078124	0.041684	0.017294	0.999604
	0.6	Team 5	0.100386	0.310491	0.177306	0.10865	0.121905	0.091505	0.06114	0.02808	0.00033	0	0.999793
		Team 6	0	0	0.062565	0.313537	0.374157	0.086692	0.088482	0.046332	0.027972	0	0.999737
		Team 10	0.1823	0.02908	0.05623	0.05138	0.04701	0.03734	0.05249	0.062287	0.050887	0.430987	0.99999

Bayesian projections for competition **Bins 1-10**

Team 1

	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.96408	0.03592	0	0	0	0	0	0	0	0	1
Team 2	0.98297	0	0.000807	0	0.000484	0	0	0	0	0.015739	1
Team 3	0.990399	0	0.001695	0.000541	0.001663	0.004823	0	0	0	0.00088	1
Team 5	0.968045	0.031955	0	0	0	0	0	0	0	0	1
Team 6	0.996921	0	0.002396	0	0.000269	0	0	0	0	0.000414	1
Team 10	0.962174	0	0	0	0	0.013487	0.001381	0.003053	0	0.019905	1

Team 3 **Bins 1-10**

	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.970772	0.029228	0	0	0	0	0	0	0	0	1
Team 2	0	0	0.015866	0	0.213101	0	0	0	0	0.771033	1
Team 3	0.998524	0	0.00073	3.55E-05	0.000142	0.000544	0	0	0	2.46E-05	1
Team 5	0.990013	0.009987	0	0	0	0	0	0	0	0	1
Team 6	0	0	0.163568	0	0.725182	0	0	0	0	0.111251	1
Team 10	0	0	0	0	0	0.634903	0.061382	0.151531	0	0.152183	1

Team 5 **Bins 1-10**

	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.999383	0.000617	0	0	0	0	0	0	0	0	1
Team 2	0.99723	0	6.26E-05	0	0.000109	0	0	0	0	0.002599	1
Team 3	0.770326	0	0.004148	0.002243	0.021125	0.201127	0	0	0	0.001032	1
Team 5	0.999931	6.91E-05	0	0	0	0	0	0	0	0	1
Team 6	0.999958	0	1.26E-05	0	2.19E-05	0	0	0	0	7.42E-06	1
Team 10	0.99686	0	0	0	0	0.001982	0.00016	0.000617	0	0.00038	1

Team 6

	Bins 1-10										
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.925679	0.074321	0	0	0	0	0	0	0	0	1
Team 2	0.990541	0	0.001783	0	0.002047	0	0	0	0	0.00563	1
Team 3	0.981861	0	0.001473	0.000297	0.001985	0.01373	0	0	0	0.000653	1
Team 5	0.989091	0.010909	0	0	0	0	0	0	0	0	1
Team 6	0.976992	0	0.019831	0	0.002146	0	0	0	0	0.001031	1
Team 10	0.989162	0	0	0	0	0.001788	0.000131	0.00044	0	0.008479	1

Team 10

	Bins 1-10										
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.966066	0.033934	0	0	0	0	0	0	0	0	1
Team 2	0	0	0.004519	0	0.525114	0	0	0	0	0.470368	1
Team 3	0	0	0.002649	0.001121	0.572375	0.397942	0	0	0	0.025912	1
Team 5	0.989307	0.010693	0	0	0	0	0	0	0	0	1
Team 6	0	0	0	0	0.781963	0	0	0	0	0.218037	1
Team 10	0.976153	0	0	0	0	0.011202	0.001669	0.003908	0	0.007068	1

Competition results (participants):

C (control)	Bins 1-10										
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0	0	0	0	0	0	0	0	0	1	1
Team 2	0	0	0	0	0	0	0	0	0	1	1
Team 3	0	0	0	0.8	0	0	0	0	0	0.2	1
Team 5	0.8	0	0	0	0.14	0	0	0	0	0.06	1
Team 6	0	0	0	0.8	0	0	0	0	0	0.2	1
Team 10	0	0	0	0	0	0.8	0	0	0	0.2	1

CS1 (waypoints--control/score)	Bins 1-10										
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0	0	0	0	0.3	0	0	0.7	0	0	1
Team 2	0	0	0	0	0.7	0	0	0	0	0.3	1
Team 3	0	0.3	0.7	0	0	0	0	0	0	0	1
Team 5	1	0	0	0	0	0	0	0	0	0	1
Team 6	0	0.3	0	0	0.7	0	0	0	0	0	1
Team 10	0	0	0	0	0.7	0	0	0	0	0.3	1

CS2 (scoring--control/score)	Bins 1-10										
	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.4	0	0	0	0.35	0.1	0	0.15	0	0	1
Team 2	0.4	0	0	0.15	0.05	0	0	0	0	0.4	1
Team 3	0	0.7	0.15	0	0.15	0	0	0	0	0	1
Team 5	1	0	0	0	0	0	0	0	0	0	1
Team 6	0	0.3	0.1	0	0.05	0.4	0	0.15	0	0	1
Team 10	0.4	0	0	0	0.2	0	0	0	0	0.4	1

S (score) **Bins 1-10**

	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.1	0	0	0	0.07	0.65	0	0	0	0.18	1
Team 2	0.25	0	0	0	0	0	0	0	0	0.75	1
Team 3	0.8	0	0.1	0	0.1	0	0	0	0	0	1
Team 5	0.9	0	0.03	0	0	0	0	0	0	0.07	1
Team 6	0.72	0	0.1	0.15	0.03	0	0	0	0	0	1
Team 10	0.1	0	0	0	0.07	0.15	0.03	0	0	0.65	1

CS--Control/scoring **Bins 1-10**

	1	2	3	4	5	6	7	8	9	10	check sum
Team 1	0.28	0	0	0	0.335	0.07	0	0.315	0	0	1
Team 2	0.28	0	0	0.105	0.245	0	0	0	0	0.37	1
Team 3	0	0.58	0.315	0	0.105	0	0	0	0	0	1
Team 5	1	0	0	0	0	0	0	0	0	0	1
Team 6	0	0.3	0.07	0	0.245	0.28	0	0.105	0	0	1
Team 10	0.28	0	0	0	0.35	0	0	0	0	0.37	1

Goals / weighting | **Complete goal rollup** | **Bins 1-10**

Control

0.1

Control/score			**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	check sum
	0.3	Team 1	0.144	0	0	0	0.1425	0.411	0	0.0945	0	0.208	1
		Team 2	0.234	0	0	0.0315	0.0735	0	0	0	0	0.661	1
Score		Team 3	0.48	0.174	0.1545	0.08	0.0915	0	0	0	0	0.02	1
	0.6	Team 5	0.92	0	0.018	0	0.014	0	0	0	0	0.048	1
		Team 6	0.432	0.09	0.081	0.17	0.0915	0.084	0	0.0315	0	0.02	1
		Team 10	0.144	0	0	0	0.147	0.17	0.018	0	0	0.521	1

difference squared between actual results and bayesian projections for competition
far right gives rms error

Team 1					Bins 1-10					
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10rms error**
Team 1	0.673	0.001	0.000	0.000	0.020	0.169	0.000	0.009	0.000	0.043	0.303
Team 2	0.561	0.000	0.000	0.001	0.005	0.000	0.000	0.000	0.000	0.416	0.314
Team 3	0.261	0.030	0.023	0.006	0.008	0.000	0.000	0.000	0.000	0.000	0.181
Team 5	0.002	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.025
Team 6	0.319	0.008	0.006	0.029	0.008	0.007	0.000	0.001	0.000	0.000	0.195
Team 10	0.669	0.000	0.000	0.000	0.022	0.024	0.000	0.000	0.000	0.251	0.311

Team 3					Bins 1-10					
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10rms error**
Team 1	0.684	0.001	0.000	0.000	0.020	0.169	0.000	0.009	0.000	0.043	0.304
Team 2	0.055	0.000	0.000	0.001	0.019	0.000	0.000	0.000	0.000	0.012	0.094
Team 3	0.269	0.030	0.024	0.006	0.008	0.000	0.000	0.000	0.000	0.000	0.184
Team 5	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.028
Team 6	0.187	0.008	0.007	0.029	0.402	0.007	0.000	0.001	0.000	0.008	0.255
Team 10	0.021	0.000	0.000	0.000	0.022	0.216	0.002	0.023	0.000	0.136	0.205

Team 5					Bins 1-10					
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10rms error**
Team 1	0.732	0.000	0.000	0.000	0.020	0.169	0.000	0.009	0.000	0.043	0.312
Team 2	0.583	0.000	0.000	0.001	0.005	0.000	0.000	0.000	0.000	0.433	0.320
Team 3	0.084	0.030	0.023	0.006	0.005	0.040	0.000	0.000	0.000	0.000	0.137
Team 5	0.006	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.030
Team 6	0.323	0.008	0.007	0.029	0.008	0.007	0.000	0.001	0.000	0.000	0.196
Team 10	0.727	0.000	0.000	0.000	0.022	0.028	0.000	0.000	0.000	0.271	0.324

Team 6					Bins 1-10						
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**rms error**
Team 1	0.611	0.006	0.000	0.000	0.020	0.169	0.000	0.009	0.000	0.043	0.293
Team 2	0.572	0.000	0.000	0.001	0.005	0.000	0.000	0.000	0.000	0.430	0.317
Team 3	0.252	0.030	0.023	0.006	0.008	0.000	0.000	0.000	0.000	0.000	0.179
Team 5	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.028
Team 6	0.297	0.008	0.004	0.029	0.008	0.007	0.000	0.001	0.000	0.000	0.188
Team 10	0.714	0.000	0.000	0.000	0.022	0.028	0.000	0.000	0.000	0.263	0.320

Team 10					Bins 1-10						
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**rms error**
Team 1	0.676	0.001	0.000	0.000	0.020	0.169	0.000	0.009	0.000	0.043	0.303
Team 2	0.055	0.000	0.000	0.001	0.204	0.000	0.000	0.000	0.000	0.036	0.172
Team 3	0.230	0.030	0.023	0.006	0.231	0.158	0.000	0.000	0.000	0.000	0.261
Team 5	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.028
Team 6	0.187	0.008	0.007	0.029	0.477	0.007	0.000	0.001	0.000	0.039	0.275
Team 10	0.692	0.000	0.000	0.000	0.022	0.025	0.000	0.000	0.000	0.264	0.317

difference squared between actual results and team projections for competition

far right gives rms error

Team 1

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**rms error**
					Bins 1-10						
Team 1	0.039	0.045	0.007	0.001	0.004	0.135	0.005	0.000	0.002	0.031	0.164
Team 2	0.016	0.074	0.014	0.006	0.007	0.026	0.008	0.001	0.000	0.435	0.242
Team 3	0.170	0.011	0.005	0.001	0.000	0.015	0.012	0.012	0.012	0.007	0.157
Team 5	0.843	0.000	0.000	0.005	0.011	0.054	0.122	0.026	0.001	0.002	0.326
Team 6	0.128	0.000	0.000	0.004	0.000	0.000	0.011	0.007	0.017	0.007	0.132
Team 10	0.005	0.008	0.007	0.011	0.002	0.006	0.007	0.012	0.015	0.160	0.153

Team 3

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**rms error**
					Bins 1-10						
Team 1	0.021	0.003	0.093	0.070	0.001	0.090	0.004	0.006	0.000	0.039	0.181
Team 2	0.052	0.005	0.022	0.084	0.004	0.005	0.017	0.004	0.001	0.421	0.248
Team 3	0.230	0.029	0.012	0.000	0.000	0.008	0.007	0.031	0.090	0.008	0.204
Team 5	0.559	0.218	0.024	0.005	0.001	0.002	0.000	0.000	0.000	0.001	0.285
Team 6	0.187	0.008	0.016	0.011	0.017	0.002	0.014	0.000	0.000	0.000	0.160
Team 10	0.021	0.000	0.000	0.000	0.009	0.020	0.107	0.050	0.002	0.259	0.216

Team 5

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**rms error**
					Bins 1-10						
Team 1	0.004	0.061	0.010	0.000	0.019	0.120	0.007	0.006	0.031	0.020	0.167
Team 2	0.000	0.028	0.009	0.000	0.002	0.004	0.017	0.011	0.009	0.368	0.212
Team 3	0.190	0.024	0.021	0.031	0.045	0.028	0.001	0.000	0.009	0.001	0.187
Team 5	0.617	0.129	0.085	0.018	0.000	0.000	0.000	0.000	0.000	0.001	0.292
Team 6	0.019	0.005	0.004	0.020	0.001	0.000	0.004	0.002	0.021	0.004	0.090
Team 10	0.021	0.000	0.000	0.000	0.022	0.029	0.000	0.139	0.137	0.076	0.206

Team 6					Bins 1-10						
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**rms error**
Team 1	0.021	0.002	0.003	0.000	0.001	0.076	0.051	0.023	0.006	0.033	0.147
Team 2	0.012	0.129	0.020	0.000	0.002	0.000	0.001	0.000	0.004	0.232	0.200
Team 3	0.196	0.017	0.009	0.000	0.002	0.074	0.029	0.034	0.008	0.000	0.192
Team 5	0.534	0.115	0.054	0.012	0.002	0.001	0.000	0.000	0.000	0.001	0.268
Team 6	0.000	0.007	0.004	0.019	0.001	0.001	0.004	0.000	0.000	0.069	0.102
Team 10	0.004	0.000	0.000	0.000	0.016	0.026	0.000	0.012	0.112	0.066	0.154

Team 10					Bins 1-10						
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**rms error**
Team 1	0.063	0.000	0.007	0.014	0.001	0.083	0.005	0.006	0.000	0.043	0.149
Team 2	0.023	0.008	0.009	0.005	0.007	0.003	0.008	0.023	0.021	0.392	0.223
Team 3	0.217	0.026	0.011	0.038	0.055	0.013	0.005	0.006	0.002	0.000	0.193
Team 5	0.672	0.096	0.025	0.012	0.012	0.008	0.004	0.001	0.000	0.002	0.288
Team 6	0.187	0.008	0.000	0.021	0.080	0.000	0.008	0.000	0.001	0.000	0.175
Team 10	0.001	0.001	0.003	0.003	0.010	0.018	0.001	0.004	0.003	0.008	0.072

Appendix F: Expected bin ranking of teams for Exhibition 1 and 2 and the Competition

Far left column indicates the team whose evaluations produced the Exhibition 1, Exhibition 2 and Competition rankings as a function of expected bin in which a team should be found, as evaluated by each of the evaluating teams (i.e., Teams 1, 3, 5, 6, and 10). The teams are ranked highest to lowest with the highest at the top and lowest at the bottom. To the right of the Exhibition 1, Exhibition 2 and Competition rankings are the results, in the form of team rankings also, for each of the respective events. To the right of the results of Exhibition 1 and Exhibition 2 are the respective posteriors projecting the success of the teams in the subsequent events, Exhibition 2 and the Competition, respectively.

		Rankings for Exhibition 1	Results Exhibition 1	Posterior for Exhibition 2	Rankings for Exhibition 2	Results Exhibition 2	Posterior for Competition	Rankings for Competition	Results Competition
Team 1	rank	team		team	team	team	team	team	team
evaluation of	1	2	assumed average	3	1	10	10	5	2
teams produces rank	2	5	performance for	2	3	3	2	10	10
(best to worst, shown	3	10	all teams	5	6	2	3	3	1
top to bottom) for	4	3		10	10	6	1	6	6
Ex. 1, Ex. 2, and	5	6		6	5	1,5	5	2	3
Competition	6	1		1	2		6	1	5
Team 3	rank	team		team	team	team	team	team	team
evaluation of	1	10	assumed average	10	10	10	2	3	2
teams produces rank	2	6	performance for	2	2	3	10	10	10
(best to worst, shown	3	2	all teams	6	6	2	6	2	1
top to bottom) for	4	1		1	1	6	1	6	6
Ex. 1, Ex. 2, and	5	3		3	5	1,5	5	1	3
Competition	6	5		5	3		3	5	5

Team 5	rank	team		team	team	team	team	team	team
evaluation of	1	3	assumed average	3	3	10	3	10	2
teams produces rank	2	2	performance for	2	2	3	2	1	10
(best to worst, shown	3	5	all teams	5	1	2	10	3	1
top to bottom) for	4	6		10	10	6	1	6	6
Ex. 1, Ex. 2, and	5	10		6	6	1,5	6	2	3
Competition	6	1		1	5		5	5	5
Team 6	rank	team ****		team	team	team	team	team	team
evaluation of	1	5	assumed average	1	5	10	10	10	2
teams produces rank	2	1	performance for	6	1	3	3	1	10
(best to worst, shown	3	3	all teams	3	3	2	1	3	1
top to bottom) for	4	2		5	2	6	2	6	6
Ex. 1, Ex. 2, and	5	6		2	6	1,5	6	2	3
Competition	6	10		10	10		5	5	5
****	Team 6 eval	uations used	From Exhibition 2	since none	provided from	Exhibition 1.			
Team 10	rank	team		team	team	team	team	team	team
evaluation of	1	6	assumed average	3	2	10	2	10	2
teams produces rank	2	2	performance for	6	10	3	6	2	10
(best to worst, shown	3	3	all teams	2	3	2	3	3	1
top to bottom) for	4	10		1	5	6	10	6	6
Ex. 1, Ex. 2, and	5	5		5	1	1,5	1	5	3
Competition	6	1		10	6		5	1	5

intentionally left blank

Appendix G: Each team's methodology in assigning scores

In Section 9.1.1 a synopsis of the methods used by the teams in assigning scores is offered. A description of each of the team-specific methods is found below.

Method used by Team 1.

1. Read the Wiki journal daily to learn the best practices of other teams and avoid problems encountered by other teams. The Wiki Journal, found on the MASLAB website [1], is a journal in which the design teams update their daily progress.
2. Observe peers in the lab and make comparisons to own design and progress.
3. Keep track of how much effort is put into the robot as an indication of expected success in the upcoming contests.
4. Use performance in previous events as evidence in evaluations of projected success in future events.

Method used by Team 3.

1. Observe peers in the lab environment.
2. Speak with the other robot design teams to identify their best practices and problems encountered.
3. Use performance in previous events as evidence in evaluations of projected success in future events.

Method used by Teams 5 and 6.

1. Observe peers in the lab environment.

2. Use performance in previous events as evidence in evaluations of projected success in future events.

Method used by Team 10.

1. Observe peers in the lab environment. Question whether other teams often lack direction or consistently know what they are doing? Ask whether teams are able to build sensors on their own? If they can, whether they are performing well above average and are expected to do well in the contests. Ask whether teams are using water-jetting (i.e., a sophisticated manufacturing method) for the robot chassis? If a team is using such a method in the manufacture of the robot chassis, their robot is likely to be fairly sophisticated overall and excellent performance thus is expected. Ask whether teams are consistently making last minute preparations prior to the contests? If so, then poor performance will likely result.

2. Use performance in previous events as evidence in evaluations of projected success in future events.

Appendix H: Making the evaluation process easier for each team

In Section 9.1.2 a synopsis of the what would have made the evaluation process easier is offered. A description of what would have made the evaluation process easier for each of the teams is found below.

Team 1

The capabilities of the robots were not equal in each of the contests (i.e., full capability was not achieved until the final contest, the Competition). This fact meant that some of the questions asked in the questionnaire did not apply until Exhibition 2 and possibly until the Competition. This fact meant that the questionnaires needed to be better suited to each event such that not only a current but also a fairly exact assessment of each robot capability could be obtained by filling out such surveys.

Team 3

Nothing would have made the evaluation process easier. The use of surveys was a good way to evaluate a robot's overall capability.

Team 5

More time should be spent in filling out the questionnaires. This was, however, a difficult issue to resolve as preparation of the robot for each of the three events was a highly time constrained process.

Team 6

Be consistent in filling out the questionnaires. That is, having previous questionnaire(s) available could assist in maintaining greater consistency in answering the questionnaires.

Team 10

Completing the first evaluation was difficult. Since the contests themselves offer the best evidence of a robot's capability, the first evaluation was based largely on observations made in the lab prior to Exhibition 1. This meant that more time must be spent in making not only the first evaluation but also the follow-on evaluations. The additional time could have been spent in discussing problems as well as best practices and plans of attack with each of the competing teams and in reading the Wiki journal [1].